Bilingualism: Language and Cognition

cambridge.org/bil

The reduced embodiment of a second language

Tal Norman[1] and Orna Peleg[2]

[1]Department of Communication Sciences and Disorders University of Haifa, Haifa, Israel and [2]The program of cognitive studies of language use and Sagol school of neuroscience Tel Aviv University, Tel Aviv, Israel

Research Article

Cite this article: Norman T, Peleg O (2022). The reduced embodiment of a second language. *Bilingualism: Language and Cognition* **25**, 406–416. https://doi.org/10.1017/S1366728921001115

Received: 27 January 2021
Revised: 14 November 2021
Accepted: 19 November 2021
First published online: 14 December 2021

Keywords:
L2 sentence comprehension; L1-L2 processing differences; Cross-language influences; Embodied cognition; Sentence picture verification task

Address for correspondence:
Tal Norman, Department of Communication Sciences and Disorders University of Haifa, 199 Abba Khoushi Ave., 3498838, Haifa, Israel
talnorm@gmail.com

Abstract

Substantial evidence indicates that first language (L1) comprehension involves embodied visual simulations. The present study tested the assumption that a formally learned second language (L2), which is less related to real-life experiences, is processed in a less embodied manner relative to a naturally acquired L1. To this end, bilingual participants completed the same task in their L1 and L2. In the task, they read sentences and decided immediately after each sentence whether a pictured object had been mentioned in the preceding sentence. Responses were significantly faster when the shape of the object in the picture matched rather than mismatched the sentence-implied shape, but only in the L1, and only when the L1 block was performed before the L2 block. These findings suggest that embodied visual simulations are reduced in a formally learned L2 and may be subjected to cross-language influences.

1. Introduction

Theories of embodied cognition postulate that conceptual knowledge is grounded in sensorimotor mechanisms. Accordingly, language comprehension involves not only the activation of linguistic representations, but also the activation of different types of non-verbal modal representations (e.g., visual, affective, motor) associated with the stated events (e.g., Barsalou, 1999, 2008; Zwaan, 2004). In particular, Barsalou, Santos, Simmons and Wilson (2008) proposed that two distinct, but interconnected systems are involved in conceptual processing – a linguistic system that uses word association to represent meaning, and a simulation system that uses non-verbal sensorimotor knowledge to reenact real life experiences – resulting in both linguistic-based and simulation-based comprehension processes.

In line with these theories, many studies have shown that during sentence processing, first language (L1) comprehenders spontaneously activate perceptual visual information that corresponds to the visual characteristics of verbally described objects (e.g., De Koning, Wassenburg, Bos & Van der Schoot, 2017a; Sato, Schafer & Bergen, 2013; Stanfield & Zwaan, 2001; Winter & Bergen, 2012; Yaxley & Zwaan, 2007; Zwaan, Madden, Yaxley & Aveyard, 2004; Zwaan, Stanfield & Yaxley, 2002). These studies have used a sentence picture verification (SPV) task, in which participants typically read (or listen to) sentences (e.g., "The ranger saw the eagle in the sky") and are asked to verify whether or not a pictured object, presented immediately after each sentence, had been mentioned in the sentence. Importantly, the visual characteristics (e.g., shape) of the object in the sentence (e.g., eagle) are determined by its described location (e.g., sky – an eagle with its wings spread out) and are not explicitly stated in the sentence. In all critical trials, the pictured object is mentioned in the sentence – however, its visual features can either match (e.g., a picture of an eagle with its wings spread out) or mismatch (e.g., a picture of an eagle with its wings folded) the sentence-implied features. Participants usually respond faster when the visual features of the pictured object match, rather than mismatch, the sentence-implied features. This visual effect indicates that L1 comprehenders routinely simulate visual characteristics of the linguistic content (Zwaan & Madden, 2005).

Indeed, substantial evidence supports an embodied view of language comprehension (for a review see Barsalou, 2008) – however, the majority of findings come from research on L1 processing. Crucially, the embodiment of a second language (L2) is relatively unexplored, and existing L2 research has focused mainly on the activation of motor and affective knowledge (e.g., Hsu, Jacobs & Conrad, 2015; Vukovic & Shtyrov, 2014). Thus, it is still unclear whether the comprehension of an L2 involves the construction of perceptual visual simulations, and if so, whether these simulations are activated to the same extent as in the L1. Therefore, to further test the relationship between visual simulations and L2 comprehension, the present study focused on late bilinguals, who have learned and used their L2 in a formal manner, outside of the environment where it is naturally and constantly spoken. It was assumed that, among this type of bilinguals, one of the fundamental differences between L1 and L2 comprehension resides in their ability to construct rich and detailed mental simulations, because their L2 usage is far less accompanied by real-life experiences.

© The Author(s), 2021. Published by Cambridge University Press

CAMBRIDGE UNIVERSITY PRESS

Embodied theories further assume that neural networks associated with the processing of real-life experiences and those associated with language processing become strongly linked over time, because they are frequently activated together during language acquisition and use (Pulvermüller, 1999, 2013; Zwaan & Madden, 2005). Thus, the coupling between sensorimotor and linguistic experiences may play a critical role in shaping embodied language comprehension (e.g., Dekker, Mareschal, Johnson & Sereno, 2014; Holt & Beilock, 2006). Under this assumption, experience-based differences in the acquisition and use of the L1 and the L2 may have significant consequences on how each language is linked to sensorimotor knowledge, both qualitatively, by determining the specific embodied representations evoked by each language (Jared, Poh & Paivio, 2013), and quantitatively, by determining the extent to which modality-specific representations are activated by each language (Vukovic & Shtyrov, 2014).

Naturally, babies, infants, and children learn the meaning of linguistic structures in their L1 while using their body and interacting with their environment. As a result, the conceptual representations evoked by their L1 are shaped by their physical experiences (Pexman, 2019). Growing up, they keep using their native language and experiencing the world simultaneously and constantly, and thus, their neural networks associated with L1 processing and those associated with the processing of real-life experiences become strongly linked over time. Therefore, as mentioned above, L1 users routinely simulate perceptual (e.g., Brunyé, Ditman, Mahoney, Walters & Taylor, 2010; Zwaan et al., 2002), motor (e.g., Glenberg & Kaschak, 2002; Myung, Blumstein & Sedivy, 2006; Scorolli & Borghi, 2007), and affective (e.g., Havas, Glenberg & Rinck, 2007; Niedenthal, Winkielman, Mondillon & Vermeulen, 2009) knowledge during language comprehension (but see Bedny & Caramazza, 2011; and Mahon, 2015, for an alternative view; and also Morey, Kaschak, Díez-Álamo, Glenberg, Zwaan, Lakens, Ibáñez, García, Gianelli, Jones, Madden, Alifano, Bergen, Bloxsom, Bub, Cai, Chartier, Chatterjee, Conwell, Cook, Davis, Evers, Girard, Harter, Hartung, Herrera, Huettig, Humphries, Juanchich, Kühne, Lu, Lynes, Masson, Ostarek, Pessers, Reglin, Steegen, Thiessen, Thomas, Trott, Vandekerckhove, Vanpaemel, Vlachou, Williams & Ziv-Crispel, 2021, for a failure to replicate the action-sentence compatibility effect first reported by Glenberg & Kaschak, 2002).

Conversely, late bilinguals living in their L1 environment many times acquire their L2 in a formal and less natural setting, and then use their L2 in relatively limited contexts. Thus, the acquisition and use of their L2 are far less related to the physical world and to bodily experiences, in comparison to their L1. Therefore, it is possible that the links between sensorimotor representations and L2 representations, in this type of bilinguals, are much weaker. Crucially, under such circumstances, the L2 may be less embodied relative to the L1.

Such a reduced ability to mentally simulate multimodal aspects of the L2 content may explain the foreign language effect, according to which decision-making in the L2 is not influenced by emotional biases or by visual imagery, as opposed to decision-making in the L1 (Hayakawa & Keysar, 2018; Keysar, Hayakawa & An, 2012). Thus, there are good reasons to assume that the activation of perceptual visual knowledge during language comprehension would be restricted in a late and formally learned L2, in comparison to a naturally learned L1. However, this assumption still needs to be confirmed by further empirical evidence, since only a few studies to date have investigated visual simulations during L2 comprehension, either demonstrating activation of visual knowledge in both the L1 and the L2 (Ahn & Jiang, 2018; Koster, Cadierno & Chiarandini, 2018), or showing visual effects only in the L1 (Chen, Wang, Zhang & Liu, 2020).

For example, Koster et al. (2018) demonstrated a similar pattern of visual activations during sentence reading in L1 users and in L2 learners. They used the SPV task with sentences that explicitly indicated a visual property of the object – either size (e.g., "Anna puts the lipstick {big} on the cutting board") or orientation (e.g., "Anna stands {vertically} the lipstick on the cutting board") – and tested L1 users and L2 learners (i.e., beginner, intermediate, advanced) of Spanish and German. Responses to target pictures were significantly faster on match than on mismatch trials, but only for size sentences, indicating that size information is more likely to be activated than information about spatial orientation. Additionally, the interaction between trial type (i.e., match vs. mismatch) and language type (i.e., L1, L2 beginners, L2 intermediate, L2 advanced) was not significant, suggesting that readers activate stated visual information during sentence comprehension, irrespective of language type or proficiency.

Furthermore, two studies that examined the activation of the implied shape of mentioned objects during sentence reading presented conflicting results. On the one hand, Ahn and Jiang (2018) found that L2 comprehenders, like L1 comprehenders, simulate the specific shape of verbally described objects. They compared the performance of L1-Korean users and proficient late L2-Korean users on the typical SPV task and found that responses were faster on match than on mismatch trials, in both language groups, suggesting that L1 and L2 comprehenders equally activate implied shape information about objects described in sentences.

On the other hand, Chen et al. (2020) demonstrated that the L2 and the L3 of trilinguals (i.e., L1-Cantonese, L2-Mandarin, L3-English) are less linked to perceptual visual knowledge, in comparison to the L1. They used a delayed SPV task that consisted of two phases. In the study phase, participants listened to L1, L2, and L3 sentences, presented in three separate and consecutive language blocks, and had to decide whether or not each sentence was meaningful. After a 10-min delay, in the test phase, they saw pictures of objects and had to decide whether or not each object had been mentioned in any of the sentences that were presented in the previous study phase. They found a significant interaction between shape condition (i.e., match vs. mismatch) and language type (i.e., L1 vs. L2 vs. L3), in which the shape effect was significant for L1 sentences but not for the L2 or L3 sentences. These findings indicate that trilingual readers activated and maintained shape information about mentioned objects only for L1 sentences. Furthermore, they show that while acquisition style (i.e., natural L1 learning vs. formal L2 and L3 learning) had a significant impact on the exhibited shape effect, proficiency level in the non-native languages (high-proficiency in the L2 vs. low-proficiency in the L3) did not modulate the effect.

In sum, as opposed to the large body of evidence supporting an embodied view of language comprehension in the case of an L1, and specifically, indicating that L1 comprehension involves the construction of perceptual visual simulations, a relatively small number of studies have investigated embodiment effects during L2 comprehension, yielding conflicting evidence (e.g., Dudschig, de la Vega & Kaup, 2014; Vukovic & Shtyrov, 2014). Crucially, of these, only a few studies have investigated visual effects.

Thus, to shed light on the ongoing controversy regarding the embodiment of the L2, particularly with respect to perceptual visual simulations, here we evaluated the extent to which L2

https://doi.org/10.1017/S1366728921001115 Published online by Cambridge University Press

comprehension involves visual simulations, relative to L1 comprehension, while (a) examining the activation of a highly intrinsic visual feature (i.e., shape), since it was demonstrated that intrinsic visual properties are more strongly activated than extrinsic ones (e.g., De Koning, Wassenburg, Bos & van der Schoot, 2017b; Koster et al., 2018; Zwaan & Pecher, 2012); (b) focusing on highly homogenous group of late bilinguals with similar linguistic and cultural background – native Hebrew speakers who learned English in formal settings in Israel, have lived there since birth, and achieved a high level of L2 proficiency; and (c) testing L1 and L2 processing within the same bilinguals.

2. The present study

This study examined whether visual shape features are activated to the same extent during L1 and L2 sentence reading. To this end, the typical SPV task[1] was used (e.g., Zwaan et al., 2002), in which participants were asked to read sentences and then respond to target pictures. All sentences described an object (e.g., lemon) in a specific location (e.g., garden) implying its shape (e.g., whole lemon). After each sentence (e.g., "The girl saw the lemon in the garden), a picture of an object (e.g., lemon) was presented and participants decided whether or not the pictured object had been mentioned in the sentence. On critical trials, the pictured object was indeed mentioned in the sentence. However, in the match condition its shape was similar to the shape implied by the sentence (e.g., a picture of a whole lemon), whereas in the mismatch condition its shape was different (e.g., a picture of a lemon slice). If readers mentally simulate the described situation, and thus, strongly activate visual shape information during sentence comprehension, they should exhibit a significant shape effect (i.e., faster responses in the match than in the mismatch condition). However, if readers rely mainly on linguistic mechanisms, and thus, do not construct visual simulations, then their response latencies in the match and the mismatch conditions should not differ.

To test whether L2 sentence reading involves visual shape simulations, and if so, whether these simulations are activated to the same extent as in the L1, late Hebrew–English bilinguals performed the task in both their L1-Hebrew and their L2-English. Thus, differences between the two languages in the activation of implied shape information were revealed by comparing the shape effect exhibited in the L1-task and in the L2-task. In line with previous L1 studies (e.g., Peleg, Ozer, Norman & Segal, 2018; Zwaan et al., 2002), participants were expected to demonstrate a significant shape effect in their L1. Namely, L1 sentences were predicted to substantially activate implied shape information. In addition, the L2 of this group of bilinguals was expected to produce weaker visual shape activations, relative to the L1, due to its formal manner of acquisition and use (Chen et al., 2020). Thus, the shape effect was predicted to be significantly reduced in the L2-task, relative to the L1-task.

[1]The SPV task (Zwaan et al., 2002) was chosen because, as opposed to other simulation-inducing paradigms (e.g., Morey et al., 2021), this paradigm has been well-established in terms of producing the visual shape effect in the L1. Indeed, this effect has been replicated in numerous L1 studies (e.g., De Koning, Wassenburg, Bos & van der Schoot, 2017b; Engelen, Bouwmeester, De Bruin & Zwaan, 2011; Hirschfeld, Zwitserlood & Dobel, 2011), including studies examining L1-Hebrew (Peleg et al., 2018) and L1-English (e.g., Zwaan & Madden, 2006; Zwaan & Pecher, 2012), the two languages that were tested in the current study, ruling out the possibility that the shape effect would be modulated by the specific language (Hebrew/English) rather than by the nativeness of language (L1/L2) as was intended.

2.1. Method

Participants

The participants were 80 students from Tel Aviv University (30 males, 50 females; ages 18-30, mean = 25, SD = 2.51)[2]. All were late Hebrew–English bilinguals – native Hebrew speakers living in Israel (their L1 environment), who did not live in any English-speaking country over 6 months, spoke only Hebrew until the age 6, learned English in a formal school setting in Israel for 8-12 years (after the age 6), and were proficient in their L2-English (all had met the university entrance requirements in English and scored at least 120/150 in the English section of the psychometric exam)[3]. In addition, participants were right-handed based on the Edinburgh Handedness Inventory (Oldfield, 1971), free of cognitive deficits, and with normal or corrected to normal vision. The study was approved by the ethical committee of Tel-Aviv University and all participants signed an informed consent approving their participation in the study.

L2 Proficiency Measures

Participants completed a detailed language-history and self-rating questionnaire (designed based on Marian, Blumenfeld & Kaushanskaya, 2007 and Lemhöfer & Broersma, 2012) regarding their L2-English background, their English proficiency level, and their average hours-per-week current use of English. To objectively evaluate lexical knowledge in the L2-English, they performed an online lexical decision test for advanced learners of English called LexTALE (Lemhöfer & Broersma, 2012; www.lextale.com). Additionally, their English score on the psychometric exam was collected. This exam tests vocabulary knowledge as well as the ability to complete sentences appropriately, to restate sentences, and to comprehend complex texts in English on an academic level (for an example of the test see https://www.nite.org.il/files/psych/new_psych/english-eng.pdf). See Table 1 for participants' proficiency measures in the L2-English.

Materials

The critical stimuli consisted of a list of 56 objects, resulting in 56 pairs of pictures, 56 pairs of Hebrew sentences, and 56 pairs of English sentences (the exact translation of the Hebrew ones). Each pair of pictures presented the same object (e.g., lemon) in two different shapes (e.g., whole lemon vs. lemon slice). All sentences had the same structure: "The *person* saw the *object* in/on the *location*". Each sentence-pair described the same object (e.g., lemon) in two different locations (e.g., garden vs. tea), implying two different object's shapes. For example, the sentence "The girl saw the lemon in the garden" implies a shape of a whole lemon, whereas the sentence "The girl saw the lemon in the tea" implies a shape of a lemon slice.

To create the two shape conditions (match/mismatch), each one of the sentences in each pair was matched with the two pictures depicting the verbally described object in two different

[2]The sample size of 80 participants (10 participants X 8 lists) was set a priori and is in accordance with that used in previous studies, which employed the same SPV task and tested 10 participants (or less) per experimental list (e.g., Dijkstra, Yaxley, Madden & Zwaan, 2004; Madden & Dijkstra, 2009; Madden & Zwaan, 2006; Peleg et al., 2018; Zwaan et al., 2002).

[3]It is important to note that although Arabic is another language spoken in Israel, our sample included only native Hebrew speakers that mostly have no (or very limited) knowledge in Arabic. Specifically, out of 80 participants, only 8 participants reported some knowledge of Arabic. Moreover, all participants were Jewish students who live in a Hebrew-speaking environment. Thus, their exposure to other languages (e.g., Arabic) is likely to be minimal.

https://doi.org/10.1017/S1366728921001115 Published online by Cambridge University Press

Table 1. Participants' proficiency measures in the L2-English

Measure Name		Mean (SD)
Age of acquisition (in years)		7.6 (1.26)
Lexical decision test (LexTALE; max score 100)		73.00 (9.94)
Psychometric exam (max score 150)		138.11 (8.35)
Self-rated proficiency on a 7-point scale (1 = very low proficiency; 7 = very high proficiency)	Overall	5.94 (0.74)
	Reading	6.03 (0.87)
	Writing	5.44 (1.07)
	Listening	6.13 (0.74)
	Speaking	5.63 (0.89)
Self-estimated average use (hours per week)	Overall	17.09 (21.49)
	Reading	9.96 (14.67)
	Writing	5.55 (12.39)
	Listening	13.11 (18.17)
	Speaking	3.07 (9.18)

shapes. In the match condition, the shape of the pictured object and the sentence-implied shape were the same, whereas in the mismatch condition they were different. Target Language (L1-Hebrew/L2-English), Sentence Version (shape 1/shape 2) and Picture Version (shape 1/shape 2) were counterbalanced across 8 experimental lists. To avoid repetition, each participant saw only one list of 56 critical objects, which were divided into two sub-lists, one for the L1-Hebrew task and one for the L2-English task. Each sub-list comprised 28 sentence-picture combinations, including 14 combinations in the match condition and 14 combinations in the mismatch condition, all requiring a "Yes" response.

To equate the number of "Yes" and "No" responses in each sub-list, additional filler items were created. These consisted of 112 sentence-picture combinations, 56 in Hebrew and 56 in English. In each language, 14 combinations presented a picture of an object that is mentioned in the sentence ("Yes" response) and 42 combinations presented a picture of an object that is not mentioned in the sentence ("No" response). Notably, the 14 fillers that required a "Yes" response presented a picture of the described location, rather than the described object (e.g., a picture of a table was presented after the sentence "The boy saw the laptop on the *table*"). This was done in order to prevent participants from merely paying attention to the main object in the sentence (e.g., laptop). These 56 fillers in each language were added to each critical sub-list, resulting in an equal number (42) of "Yes" and "No" responses.

In sum, each sub-list (L1-Hebrew or L2-English) consisted 84 items (i.e., sentence-picture combinations): 28 critical items presenting pictures of objects that are mentioned in the sentence (14 items in each shape condition), and 56 filler items, 14 presenting pictures of objects that are mentioned in the sentence, and 42 presenting pictures of objects that are not mentioned in the sentence.

Note that the critical stimuli were designed to nullify the influence of potential differences between the two pictures of each object, other than their visual shape, by using two sentence versions for each object and presenting each one of the two pictures in both the match and the mismatch conditions. In this design, an exhibited difference between the match and the mismatch conditions, for a specific object, will be comprised of an equal number of responses to both pictures of the same object. See Table 2 for examples of critical and filler items.

Pre-tests

(1) To ensure that participants would be likely to understand the meaning of the critical English sentences, 20 students who did not participate in the main experiment translated all English sentences to Hebrew. Only sentences that were correctly translated by at least 80% of the students were included in the main experiment. (2) To ensure that all critical pictures activate their intended target word, another group of 20 students named the designated pictures in their L1-Hebrew. Only objects that both of their pictures elicited the required naming by at least 80% of the students were included in the main experiment. (3) To ensure that the critical sentences imply the intended object's shape, all sentence-pairs describing the same object were divided to form two lists of sentences. Each list consisted of only one sentence for each object. Sentences in each list were presented along with both the shape-matching and the shape-mismatching pictures. Another group of 40 students (20 per-list) were asked to choose the picture that best fit each sentence (following Connell, 2007). Only objects that both of their sentences had the picture from the match condition chosen by at least 80% of the students were used in the main experiment.

Design and procedure

To examine L1 and L2 processing within the same bilinguals all participants performed the task in both their L1-Hebrew and their L2-English. Thus, a 2X2 factorial design was used with Shape Condition (match/mismatch) and Target Language (L1-Hebrew/L2-English) as within-participant independent variable.

Participants were tested individually in a sound-attenuated room, seated in front of a computer screen with a screen-eye distance of 57 cm, and their head position was controlled using a chin-rest. They performed the L1-task and the L2-task consecutively, in two separate blocks, within the same session. The order of the two language blocks was counterbalanced across participants, such that half performed the L1-Hebrew block first and the L2-English block second, and half performed the L2-English block first and the L1-Hebrew block second. Each participant saw only one experimental list that included two different sub-lists of critical objects, one for each language task. The stimuli in each sub-list were presented in a random order, and the 8 experimental lists were counterbalanced across participants. Crucially, each participant saw each critical object only once.

The experimental session took approximately 30 minutes and consisted of 6 parts administered in a fixed order: (1) performing the handedness assessment using the computerized version (Zhang, 2014) of the Edinburgh Inventory (Oldfield, 1971); (2) filling out the L2-history and self-rating questionnaire; (3) performing the SPV task either in the L1-Hebrew or in the L2-English; (4) performing the SPV task in the other language; (5) completing an English–Hebrew translation post-test; (6) performing the online English version of the LexTALE (Lemhöfer & Broersma, 2012; http://lextale.com/takethetest.html).

Sentence Picture Verification Task

At the beginning of each language block, participants were instructed to decide as quickly and accurately as possible in

https://doi.org/10.1017/S1366728921001115 Published online by Cambridge University Press

Table 2. Examples of critical and filler items

Item	Condition	Sentence	Picture*	Response
Critical (version 1)	Match	The girl saw the lemon in the garden		Yes
	Mismatch	The girl saw the lemon in the garden		Yes
Critical (version 2)	Match	The girl saw the lemon in the tea		Yes
	Mismatch	The girl saw the lemon in the tea		Yes
Filler (location)	Related	The boy saw the laptop on the table		Yes
Filler (object)	Unrelated	The boy saw the butter in the fridge		No

Note: *Color pictures were presented in the study.

each trial whether or not the pictured object had been mentioned in the preceding sentence, by pressing a "Yes" or "No" button. Instructions were presented in Hebrew in both language blocks. After reading the instructions, participants were introduced to 4 examples of sentence-picture matching decisions, presented either in Hebrew in the L1 block, or in English in the L2 block. Then, they performed a short practice, which consisted of 6 sentence-picture combinations, presented either in Hebrew in the L1 block, or in the English in the L2 block, half requiring a "Yes" response and half requiring a "No" response. After each practice trial a visual feedback for correct and incorrect responses was provided.

All trials consisted of the same sequence of events. Initially, participants were presented with a central fixation cross for 750 ms. The offset of the cross was followed by a centrally displayed sentence presented for 4000 ms, allowing adequate processing time in both languages. Then, a central fixation cross appeared for 250 ms, followed by a centrally displayed picture that remained on the screen for 150 ms. Then, a white screen was displayed until a response was made or until 3000 ms. In each trial, response latency was measured from the onset of the picture presentation, and response accuracy was recorded.

The task was programmed and run using the E-prime software (version 10.242, Psychology Software Tools, Pittsburgh, PA). Sentences were presented in black letters on a white background, in either a Hebrew or English Times New Roman font. Pictures occupied a square in the size of 6 × 6 cm surrounded by a one cm white frame and were displayed on a gray background. RT and error data for each response were collected using a PST Serial Response Box.

English–Hebrew Translation Post-test

To ensure that participants knew the exact meaning of the critical English sentences, at the end of the experimental session, they translated to Hebrew all the critical sentences that were presented in the L2-English task.

2.2. Results

The RT data was analyzed using linear mixed effects (LME) models (Baayen, Davidson & Bates, 2008), as implemented within the 'lmer' function in the "lme4" package in R (version 3.5.2, R Core Team, 2018). The data analyses focused on the fixed effects of Shape Condition (match/mismatch) and Target Language (L1-Hebrew/L2-English) and on the interaction between them. Additionally, the fixed effects of Experimental Block (first-block/second-block) and English Proficiency Score[4] were also examined, in order to test whether responses were influenced by the block in which the task was administered or by variations between participants in English proficiency.

Three LME models were fitted to the RT data. All models included the random effects of Participants and Items (i.e., by-participants and by-item random intercepts) in order to test the major hypotheses while considering simultaneously the variance due to the random selection of participants and items. In addition, Model 1 included the fixed main effects of Shape Condition and Target Language, and the interaction between them; Model 2

[4]An English Proficiency Score was calculated for each participant by averaging the Z-scores of the 12 proficiency measures that were collected. This measure represented various subjective and objective aspects of L2 proficiency.

https://doi.org/10.1017/S1366728921001115 Published online by Cambridge University Press

included the fixed main effects of Shape Condition, Target Language, and Experimental Block, and the interactions between them; and Model 3 included the fixed main effects of Shape Condition, Target Language, Experimental Block, and English Proficiency Score, and the interactions between them.

To test whether one model provides a significantly improved fit for the data than the other two, these three models were compared using likelihood ratio tests via the 'anova' function in the "stats" package, which result in chi-square statistic. The model that best fitted the data was selected for further analysis. To evaluate the significance of the main-effects and interactions within the selected model, a type-II ANOVA with Wald Chi-square test was computed via the 'Anova' function in the "car" package. Finally, calculations of contrasts for factor interactions were performed using the Chi-Square test with the Bonferroni adjustment via the 'testInteractions' function in the "phia" package.

Data Cleanup

The entire dataset, a total of 13440 trials (4480 critical trials and 8960 filler trials), was inspected in terms of accuracy rates per-participant as well as per-item, vocabulary knowledge of critical L2-items per-participant, and RT outliers. First, accuracy rates were examined for each participant and item in each language. It was predetermined that participants and items that would have a mean accuracy rate lower than 80%, in either the L1-Hebrew or the L2-English block, would be excluded from analyses. However, none of the participants or items was rejected based on this criterion. Next, 35 English trials that were incorrectly translated in the English–Hebrew translation post-test were removed, 31 trials with RT greater than 3000 ms or smaller than 200 ms were removed, and 122 trials that fell outside the range of acceptable latencies (i.e., +/− 3.5 SD from participant's mean RT) were removed. This trimming procedure accounted for a total loss of 188 trials (1.4%). Finally, 104 critical trials (2.4%) were removed due to incorrect responses, and filler trials were excluded from the data, such that the final RT dataset consisted of correct critical trials only. Thus, 4281 data points (2169 in L1-Hebrew and 2112 in L2-English) that 80 participants produced by responding to 56 critical items were analyzed.

Data analyses

The comparison of Models 1, 2, and 3 revealed that Model 2 fitted the RT data significantly better than Model 1 ($\chi^2(4) = 105.43$, $p < .001$) and that Model 3 did not fit the data significantly better than Model 2 ($\chi^2(8) = 4.68$, $p = .79$). Indeed, the Akaike Information Criterion (AIC) of the three models (obtained from the 'summary' function in R) also revealed that Model 2 was the best fit for the data since it had the lowest AIC value (AIC: Model 1 = 55266; Model 2 = 55168; Model 3 = 55180).

Therefore, Model 2, which included the fixed main effects of Shape Condition, Target Language, and Experimental Block, the interactions between them, and the random effects of Participants and Items, was selected for further analysis. A summary of Model 2 (obtained from the 'summary' function in R) is presented in Table 3. Note that the fixed effects in Table 3 reflect simple effects relative to the reference level of the factor when the other factors are at their reference level (i.e., Shape Condition = 'match', Target Language = 'L2-English', Experimental Block = 'first-block'), rather than main effects (e.g., the coefficient for Shape Condition reflects its effect when the Target Language is 'L2-English' and the Experimental Block is 'first-block', and not across all levels). For main effects see chi-square values in the text.

Table 3. Summary of Model 2. Slopes (b), standard errors (SE), and t-values (t) of the fixed effects; Variance and standard deviation (SD) of the random effects

Fixed Effects	b	SE	t
(intercept)	646.47	20.40	31.68 ***
Shape Condition (Mismatch)	1.00	8.88	0.11
Target Language (He)	−56.94	27.71	−2.06 *
Exp. Block (Second)	−99.91	27.73	−3.60 ***
Shape Condition (Mismatch) : Target Language (He)	30.98	12.53	2.47 *
Shape Condition (Mismatch) : Exp. Block (Second)	18.64	12.63	1.48
Target Language (He) : Exp. Block (Second)	111.58	53.99	2.07 *
Shape Condition (Mismatch) : Target Language (He) : Exp. Block (Second)	−37.42	17.34	−2.11 *
Random Effects	**Variance**		**SD**
Participant (intercept)	13787		117.42
Item (intercept)	1798		42.41
Residual	21008		144.94

Sig. codes (p-values): 0 '***' 0.001 '**' 0.01 '*' 0.05 '~' 0.1 ' ' 1

Within Model 2, the main effect of Shape Condition was significant ($\chi^2(1) = 13.80$, $p < .001$), indicating that overall responses to 'match' trials (Mean = 596.29, SD = 189.51) were faster than responses to 'mismatch' trials (Mean = 612.18, SD = 195.95). In addition, the main effect of Experimental Block was significant ($\chi^2(1) = 99.04$, $p < .001$), indicating that overall responses to 'first-block' trials (Mean = 626.72, SD = 202.47) were slower than responses to 'second-block' trials (Mean = 581.64, SD = 180.01). However, the two-way interaction between Experimental Block and Target Language was only marginally significant ($\chi^2(1) = 3.04$, $p = .08$).

Importantly, the three-way interaction between Shape Condition, Target Language, and Experimental Block was significant ($\chi^2(1) = 4.45$, $p < .05$), indicating that Experimental Block modulated the interaction between Shape Condition and Target Language. Thus, the two-way interaction between Shape Condition and Target Language was analyzed separately for each Experimental Block, revealing that this interaction was significant only in the 'first-block' ($\chi^2(1) = 6.12$, $p < .05$), but not in the 'second-block' ($\chi^2(1) = .27$, $p = 1.00$).

Further examination of the effect of Shape Condition in each Target Language, separately for each Experimental Block, revealed that on 'first-block' trials, while in the 'L1-Hebrew' the shape effect was significant ($\chi^2(1) = 13.11$, $p < .01$) such that 'match' trials (Mean = 589.99, SD = 186.03) were faster than 'mismatch' trials (Mean = 622.88, SD = 205.54), in the 'L2-English', the RT-difference between 'match' trials (Mean = 646.21, SD = 208.27) and 'mismatch' trials (Mean = 648.44, SD = 204.55) was not reliable ($\chi^2(1) = .02$, $p = 1.00$). Thus, in the 'first-block', the influence of Shape Condition on speed performance was significantly evident only on 'L1-Hebrew' trials. In contrast, on 'second-block' trials, the difference between the 'match' and 'mismatch' conditions was insignificant in the L1-Hebrew ($\chi^2(1) = 2.63$, $p = .53$) as well as in the L2-English ($\chi^2(1) = 4.79$, $p = .11$), indicating that in the 'second-

https://doi.org/10.1017/S1366728921001115 Published online by Cambridge University Press

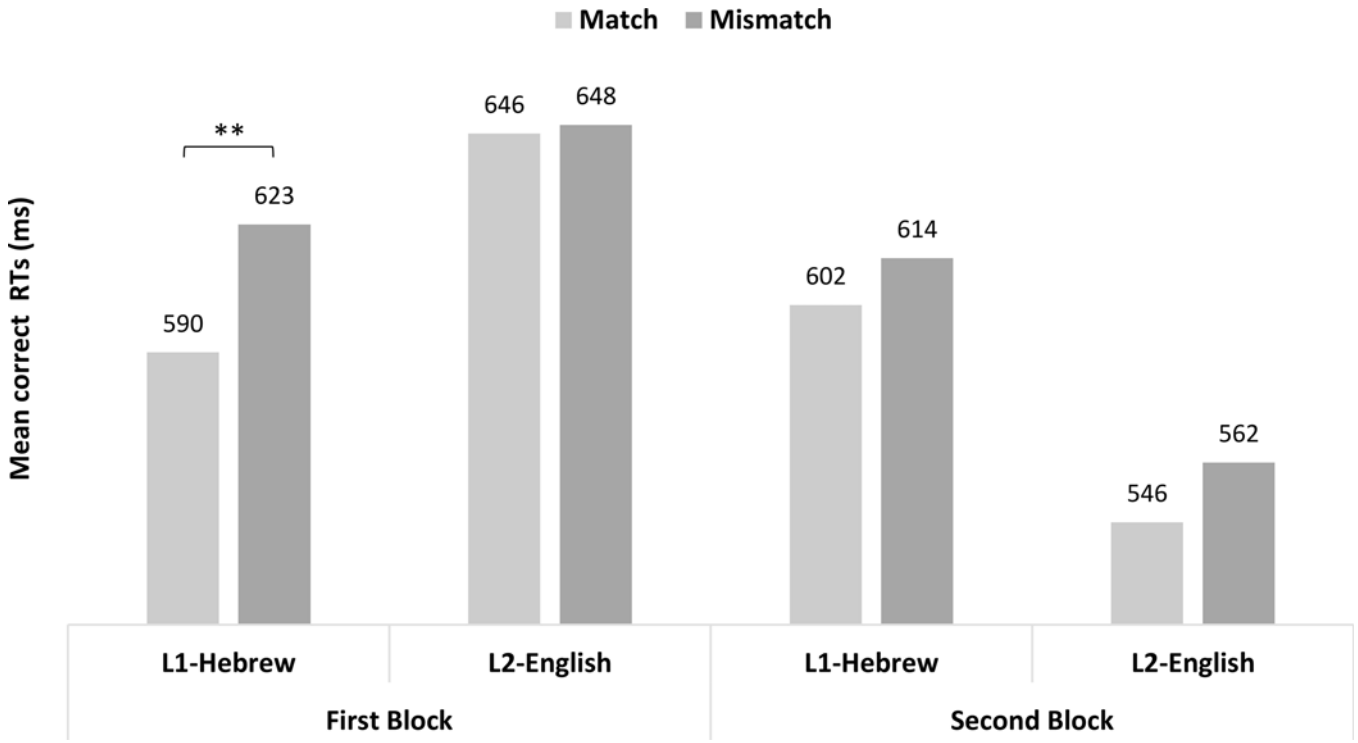


Figure 1. Mean correct RTs (ms) by Shape Condition, Target Language, and Experimental Block
Sig. codes (p-values): 0 '***' 0.001 '**' 0.01 '*' 0.05 '∼' 0.1

block', the Shape Condition had no significant influence on speed performance in both languages. Mean correct RTs (ms) by Shape Condition, Target Language, and Experimental Block, are illustrated in Figure 1.

In addition, the two-way interaction between Shape Condition and Experimental Block was analyzed separately for each Target Language, revealing that this interaction was insignificant in both the L1-Hebrew ($\chi^2(1) = 2.27$, $p = .26$) and the L2-English ($\chi^2(1) = 2.18$, $p = .28$). Thus, in both languages, the shape effect in the 'first-block' and in the 'second-block' did not significantly differ. The shape effect (ms) across Experimental Block (panel A) and Target Language (panel B) is illustrated in Figure 2.

2.3. Discussion

The current results revealed a significant shape effect only in the L1-Hebrew, and only when participants performed the L1-task before the L2-task. These findings demonstrate that L1 and L2 sentence comprehension may differ in the extent to which visual shape information is activated, such that perceptual visual activations are more likely to be generated during L1 than during L2 comprehension. Interestingly, the results further show cross-language influences, which resulted from performing the two language tasks successively, on the degree to which sentence comprehension involves perceptual visual simulations. We discuss these two findings below.

First, the fact that a significant shape effect was found only in the L1 (on first-block trials), but not in the L2, of late bilinguals that have acquired and used their L2 primarily in formal settings, supports our initial prediction that among these type of bilinguals the mental representation resulting from L2 comprehension is less grounded in sensorimotor knowledge, relative to the L1, and is consistent with previous studies, showing limited activation of perceptual (e.g., Chen et al., 2020), motor (e.g., Vukovic & Shtyrov, 2014), and affective (e.g., Hsu et al., 2015) knowledge during L2 reading.

As mentioned above, two previous studies that utilized a similar task yielded conflicting results. On the one hand, consistent with the results of the current study, Chen et al. (2020) found a significant interaction between Shape Condition (match/mismatch) and Target Language (L1/L2/L3), in which the shape effect was significant only in the L1, but not in the L2 nor in the L3, of late trilinguals. On the other hand, Ahn and Jiang (2018), observed a significant shape effect during both L1 and L2 sentence reading. This inconsistency regarding L1-L2 differences in the exhibited shape effect may be explained by differences between these studies in the experimental design, specifically in the manipulation of the language variable; and in the degree to which participants were immersed in an L2 environment at the time of testing.

First, in Ahn and Jiang's (2018) study, the language variable (L1 vs. L2) was manipulated by testing two different groups of language users (L1 vs. L2 users of Korean). Thus, the difference between L1 and L2 processing was examined in participants with different cultural and linguistic backgrounds. Additionally, within the L2-Korean group, participants varied in their L1s. This increased variance (both between the two language groups and within the L2 group) may explain Ahn and Jiang's (2018) failure to exhibit a significant difference in the shape effect between L1 and L2 sentence reading. However, in the current study (as well as in Chen et al., 2020) the language variable (L1 vs. L2) was manipulated by examining the same bilinguals in their two

https://doi.org/10.1017/S1366728921001115 Published online by Cambridge University Press

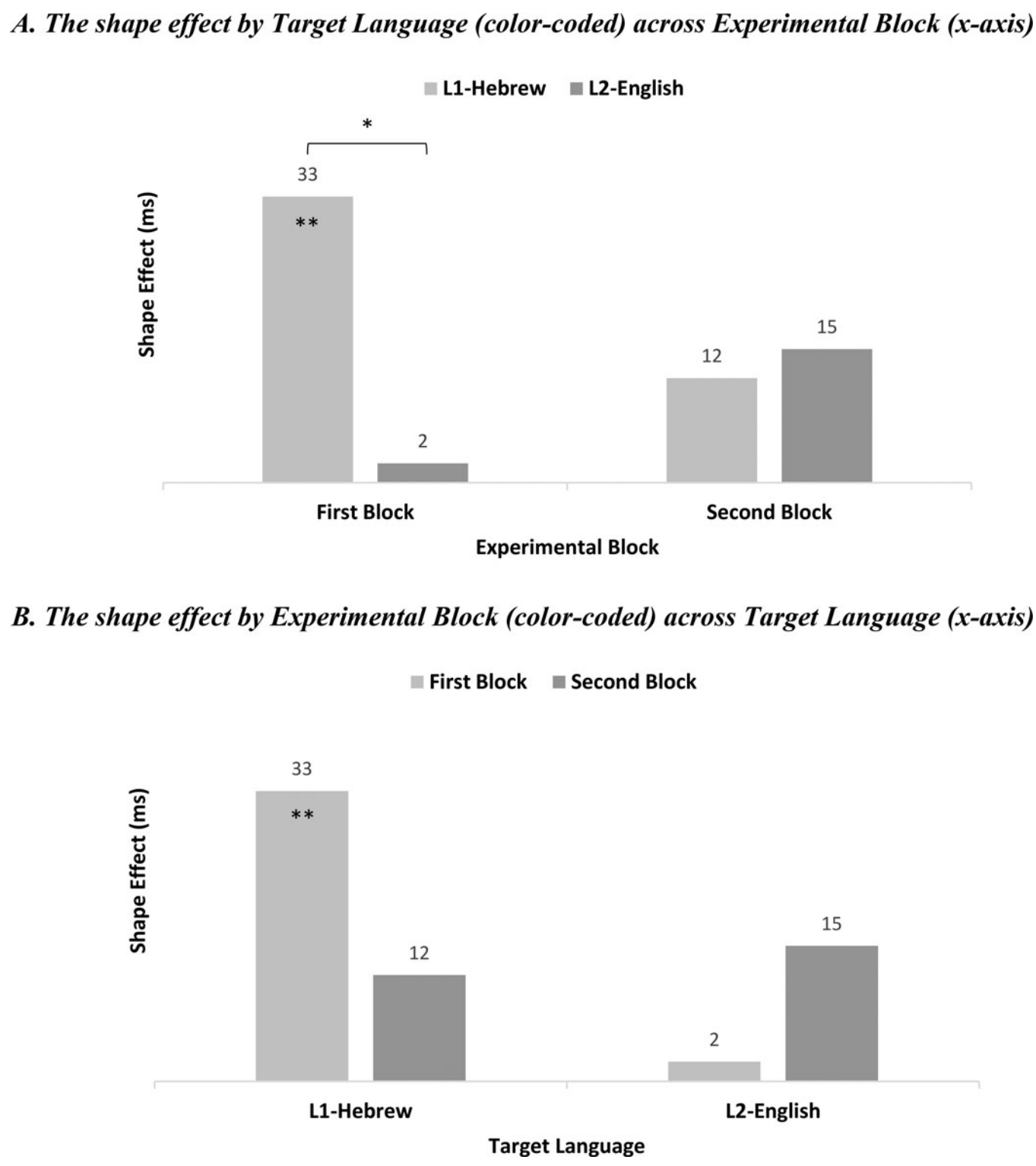


Figure 2. The shape effect* (ms) across Experimental Block and Target Language
Note: *The shape effect is the difference in mean RTs between the shape mismatch and the match conditions.
Sig. codes (p-values): 0 '***' 0.001 '**' 0.01 '*' 0.05 '~' 0.1

languages (L1-Hebrew vs. L2-English), eliminating possible differences between L1 and L2 processing in terms of participants' linguistic and cultural background.

Moreover, participants in the L2 group of Ahn and Jiang's (2018) study had lived in Korea (i.e., L2 environment) for a period of up to six years. However, in the present study, most of the participants (n = 77) had not lived in any English-speaking country (i.e., L2 environment), and the few that did (n = 3), had stayed there only for a period of up to six months. Likewise, the trilingual participants in Chen et al.'s (2020) study had lived in their L1

https://doi.org/10.1017/S1366728921001115 Published online by Cambridge University Press

environment (Cantonese) and had learned and used their L2 (Mandarin) and L3 (English) primarily in formal school settings. Therefore, it is possible that the late bilinguals tested in Ahn and Jiang's (2018) study had more immersive L2 experience that resulted in simulation-driven sentence comprehension, whereas the lack of an immersive L2 background among late bilinguals in the current study (as well as in Chen et al., 2020) resulted in sentence comprehension processes that rely mainly on linguistic mechanisms and does not involve visual simulations. Still, additional research is needed to elucidate how the learning and usage profile of bilingual population (e.g., immersed vs. un-immersed L2 learners) affect the degree of sensorimotor recruitment during L2 comprehension.

The second finding of this study is that the experimental block, in which the task was performed (first-block/second-block), modulated the L1-L2 difference in the shape effect. Specifically, when only first-block trials were analyzed, a significant interaction between Shape Condition and Target Language was observed, such that in the L1, the shape effect was significant, whereas in the L2, the match and mismatch conditions hardly differed. However, when only second-block trials were analyzed, the interaction between Shape Condition and Target Language was not evident and the shape effect was not reliable in both languages – we assume because cross-language influences on second-block trials eliminated the L1-L2 difference in the shape effect. Thus, in the L1, the shape effect was significant in the first-block but not in the second-block, probably due to the immediate recent experience with L2 sentence reading, in which comprehension relies heavily on linguistic mechanisms. By contrast, in the L2, the shape effect – although insignificant – was somewhat more considerable in the second-block relative to the first-block, probably due to the immediate recent experience with L1 sentence reading, in which comprehension involves simulating the described situation[5].

These cross-language influences are in line with previous studies reporting an effect of recent experience in the L1/L2 on task performance in the other language (e.g., Ben-Dror, Bentin & Frost, 1995; Degani, Kreiner, Ataria & Khateeb, 2019; Kreiner & Degani, 2015). For example, Ben-Dror et al. (1995) demonstrated that among Hebrew–English bilinguals, the pattern of phonological awareness during auditory word processing in the L1-Hebrew was significantly affected by whether they recently experienced word processing in their L2-English. Participants' performance in a phonological task on spoken words in the L1-Hebrew (i.e., delete the first "sound" of each word and say as fast as possible what is left of the word after this omission) was influenced by processing strategies suited for the Hebrew orthography (i.e., omitting the first two phonemes of a CVC word), when the L1-Herbew task was performed first. However, when the L1-Hebrew task was performed immediately after the L2-English task, their performance in the L1-Hebrew was more similar to their performance in the L2-English, which was influenced by processing strategies suited for the English orthography (i.e., omitting only the first phoneme of a CVC word). These results suggest that when the same task is performed in both the L1 and the L2 successively, the specific processing patterns usually employed in each language may become more similar to the processing pattern of the other language, as was also evident in the current study.

Notably, in the current study, the cross-language influence was more pronounced in the L1 than in the L2, because in the L1 the shape effect was significant in the first-block but not in the second-block, whereas in the L2 the effect was insignificant in both blocks. Thus, when only first-block trials were analyzed, in which task performance could not have been modulated by cross-language influences, the L1 results were in line with previous L1 studies which have consistently demonstrated a significant RT-facilitation when the shape of the pictured object matched, rather than mismatched, the sentence-implied shape (e.g., Peleg et al., 2018; Zwaan et al., 2002). However, when only second-block trials were analyzed, the significant shape effect that was exhibited in the L1 in the first block disappeared. These cross-language influences suggest that visual simulation processes during L1 comprehension are much more context-dependent than previously assumed.

Taken together, these findings demonstrate that simulation processes during bilingual sentence comprehension may be modulated not only by the manner of language acquisition (a naturally acquired L1 vs. a formally learned L2), but also by the presence or absence of cross-language influences (L1 before L2 vs. L1 after L2). First, we show that visual simulations are more likely to occur in the L1 than in the L2, as the visual shape effect was significantly observed only in the L1. Second, we show that even in the case of an L1, simulation effects may be context-dependent, as visual simulations were observed only when participants performed the L1-task before the L2-task. These results can be explained by embodied theories of language processing, which distinguish between comprehension processes that merely employ the linguistic system and deeper comprehension processes that employ the simulation system as well (Barsalou et al., 2008). Accordingly, an L2 that is learned formally does not establish strong links between these two systems, and thus relies more heavily on the linguistic system, as evident by the fact that in the L2-task the shape effect was insignificant in both blocks. On the other hand, a naturally learned L1 is characterized by a strong connection between the two systems that enables both types of processing – shallower processing that employs only the linguistic system, and deeper processing that includes the activation of perceptual representations in the simulation system – as evident by the fact that in the L1-task the shape effect was significant in the first-block but insignificant in the second-block.

The main limitation of the present study is related to the fact that the SPV task by itself cannot provide unequivocal evidence for the underlying mechanisms associated with the observed L1-L2 difference in the shape effect (e.g., Ostarek & Huettig, 2019). Thus, although our findings are straightforwardly explained by embodied theories that postulate a distinction between linguistic-based and simulation-based processes (Barsalou et al., 2008), more research is needed in order to confirm these preliminary conclusions. Another limitation is related to the fact that even though we recruited a highly homogenous group of participants in terms of their language background and proficiency (all native Hebrew speakers who learned English formally as an L2), we did not collect measures of exposure to other languages except the L2-English. Previous research has shown that ambient exposure to multilingualism can have an impact on language comprehension (e.g., Fan, Liberman, Keysar & Kinzler, 2015). Thus, although it is

[5]It could be that the shape effect in the second-block changed as a function of natural modification in response style over the period of the task, rather than as a function of switching to the other language. Yet, we ruled out this possibility in a follow-up analysis that tested the interaction between Trial Order and Shape Condition separately in each Experimental Block. This analysis revealed that the shape effect was not modulated by Trial Order, in both blocks, emphasizing that it was indeed the shift to the other language that altered the pattern of the shape effect in the second-block.

https://doi.org/10.1017/S1366728921001115 Published online by Cambridge University Press

unlikely that the participants in the present study were regularly exposed to other languages (i.e., they all live in a predominantly Hebrew speaking environment), future studies should include such measures.

3. Conclusions

The prevalence of non-native language users over the world (Grosjean, 2010) requires a deeper understating of the processes involved in non-native language use. Uncovering points of divergence between L1 and L2 processing may be an important step towards understating the factors crucial for mastering a non-native language. The current study revealed that the two languages of late bilinguals, who have learned and used their L2 in a formal and un-immersive manner, are processed differently, such that each language is embodied to a different extent. It appears that L1 reading is more likely to evoke perceptual visual knowledge – we assume because a native language is acquired and used in natural settings, in which linguistic information is regularly related to the physical world. In contrast, it seems that L2 reading is less likely to activate this type of embodied knowledge – we assume because, during formal learning and usage of an L2, linguistic information is far less related to non-verbal multimodal information. Future studies should examine whether simulation abilities predict comprehension abilities in both the L1 and the L2.

Acknowledgments. This work was supported by the Israeli Ministry of Education (The Zalman Aran scholarship)

References

Ahn S and Jiang N (2018) Automatic semantic integration during L2 sentential reading. *Bilingualism: Language and Cognition* **21**, 375–383.

Baayen RH, Davidson DJ and Bates DM (2008) Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language* **59**, 390–412.

Barsalou LW (1999) Perceptions of perceptual symbols. *Behavioral and brain sciences* **22**, 637–660.

Barsalou LW (2008) Grounded cognition. *Annual Review of Psychology* **59**, 617–645.

Barsalou LW, Santos A, Simmons WK and Wilson CD (2008) Language and simulation in conceptual processing. *Symbols, embodiment, and meaning*, 245–283.

Bedny M and Caramazza A (2011) Perception, action, and word meanings in the human brain: the case from action verbs. *Annals of the New York Academy of Sciences* **1224**, 81–95.

Ben-Dror I, Frost R and Bentin S (1995) Orthographic representation and phonemic segmentation in skilled readers: A cross-language comparison. *Psychological Science* **6**(3), 176–181.

Brunyé TT, Ditman T, Mahoney CR, Walters EK and Taylor HA (2010) You heard it here first: Readers mentally simulate described sounds. *Acta Psychologica* **135**, 209–215.

Chen D, Wang R, Zhang J and Liu C (2020) Perceptual Representations in L1, L2 and L3 Comprehension: Delayed Sentence–Picture Verification. *Journal of psycholinguistic research*, 1–17.

Connell L (2007) Representing object colour in language comprehension. *Cognition* **102**, 476–485.

Degani T, Kreiner H, Ataria H and Khateeb F (2019) The impact of brief exposure to the second language on native language production: Global or item specific? *Applied Psycholinguistics*, 1–31.

Dekker TM, Mareschal D, Johnson MH and Sereno MI (2014) Picturing words? Sensorimotor cortex activation for printed words in child and adult readers. *Brain and language* **139**, 58–67.

De Koning BB, Wassenburg SI, Bos LT and van der Schoot M (2017a) Size does matter: Implied object size is mentally simulated during language comprehension. *Discourse processes* **54**, 493–503.

De Koning BB, Wassenburg SI, Bos LT and van der Schoot M (2017b) Mental simulation of four visual object properties: similarities and differences as assessed by the sentence–picture verification task. *Journal of Cognitive Psychology* **29**, 420–432.

Dudschig C, de la Vega I and Kaup B (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.

Fan SP, Liberman Z, Keysar B and Kinzler KD (2015) The exposure advantage: Early exposure to a multilingual environment promotes effective communication. *Psychological science* **26**, 1090–1097.

Glenberg AM and Kaschak MP (2002) Grounding language in action. *Psychonomic bulletin & review* **9**, 558–565.

Grosjean F (2010) *Bilingual*. Harvard university press.

Havas DA, Glenberg AM and Rinck M (2007) Emotion simulation during language comprehension. *Psychonomic bulletin & review* **14**, 436–441.

Hayakawa S and Keysar B (2018) Using a foreign language reduces mental imagery. *Cognition* **173**, 8–15.

Holt LE and Beilock SL (2006) Expertise and its embodiment: Examining the impact of sensorimotor skill expertise on the representation of action-related text. *Psychonomic bulletin & review* **13**, 694–701.

Hsu CT, Jacobs AM and Conrad M (2015) Can Harry Potter still put a spell on us in a second language? An fMRI study on reading emotion-laden literature in late bilinguals. *Cortex* **63**, 282–295.

Jared D, Poh RPY and Paivio A (2013) L1 and L2 picture naming in Mandarin–English bilinguals: A test of bilingual dual coding theory. *Bilingualism: Language and Cognition* **16**, 383–396.

Keysar B, Hayakawa SL and An SG (2012) The foreign-language effect: Thinking in a foreign tongue reduces decision biases. *Psychological science* **23**, 661–668.

Koster D, Cadierno T and Chiarandini M (2018) Mental simulation of object orientation and size: A conceptual replication with second language learners. *Journal of the European Second Language Association* **2**(1).

Kreiner H and Degani T (2015) Tip-of-the-tongue in a second language: The effects of brief first-language exposure and long-term use. *Cognition* **137**, 106–114.

Lemhöfer K and Broersma M (2012) Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior research methods* **44**, 325–343. http://lextale.com/takethetest.html

Mahon BZ (2015) The burden of embodied cognition. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale* **69**, 172.

Marian V, Blumenfeld HK and Kaushanskaya M (2007) The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*.

Morey RD, Kaschak MP, Díez-Álamo AM, Glenberg AM, Zwaan RA, Lakens D, Ibáñez A, García A, Gianelli C, Jones JL, Madden J, Alifano F, Bergen B, Bloxsom NG, Bub DN, Cai ZG, Chartier CR, Chatterjee A, Conwell E, Cook SW, Davis JD, Evers E, Girard S, Harter D, Hartung F, Herrera E, Huettig F, Humphries S, Juanchich M, Kühne K, Lu S, Lynes T, Masson MEJ, Ostarek M, Pessers S, Reglin R, Steegen S, Thiessen ED, Thomas LE, Trott S, Vandekerckhove J, Vanpaemel W, Vlachou M, Williams K and Ziv-Crispel N (2021) A pre-registered, multi-lab non-replication of the action-sentence compatibility effect (ACE). *Psychonomic Bulletin & Review*.

Myung JY, Blumstein SE and Sedivy JC (2006) Playing on the typewriter, typing on the piano: manipulation knowledge of objects. *Cognition* **98**(3), 223–243.

Niedenthal PM, Winkielman P, Mondillon L and Vermeulen N (2009) Embodiment of emotion concepts. *Journal of personality and social psychology* **96**, 1120.

Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* **9**, 97–113.

Ostarek M and Huettig F (2019) Six challenges for embodiment research. *Current Directions in Psychological Science* **28**, 593–599.

https://doi.org/10.1017/S1366728921001115 Published online by Cambridge University Press

Peleg O, Ozer R, Norman T and Segal O (2018) Perceptual simulations during sentence comprehension: A comparison between typical adolescents and adolescents with autism spectrum disorder. *Journal of Neurolinguistics* **45**, 36–44.

Pexman PM (2019) The role of embodiment in conceptual development. *Language, Cognition and Neuroscience* **34**, 1274–1283.

Pulvermüller F (1999) Words in the brain's language. *Behavioral and brain sciences* **22**, 253–279.

Pulvermüller F (2013) How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics. *Trends in cognitive sciences* **17**, 458–470.

Sato M, Schafer AJ and Bergen BK (2013) One word at a time: mental representations of object shape change incrementally during sentence processing. *Language and Cognition* **5**, 345–373.

Stanfield RA and Zwaan RA (2001) The effect of implied orientation derived from verbal context on picture recognition. *Psychological science* **12**, 153–156.

Vukovic N and Shtyrov Y (2014) Cortical motor systems are involved in second-language comprehension: Evidence from rapid mu-rhythm desynchronisation. *NeuroImage* **102**, 695–703.

Winter B and Bergen B (2012) Language comprehenders represent object distance both visually and auditorily. *Language and Cognition* **4**, 1–16.

Yaxley RH and Zwaan RA (2007) Simulating visibility during language comprehension. *Cognition* **105**, 229–236.

Zhang Y (2014) Online tool for handedness assessment. Retrieved from http://zhanglab.wikicom/handedness.

Zwaan RA (2004) The immersed experiencer: Toward an embodied theory of language comprehension. *Psychology of learning and motivation* **44**, 35–62.

Zwaan RA and Madden CJ (2005) Embodied sentence comprehension. *Grounding cognition: The role of perception and action in memory, language, and thinking*, 224–245.

Zwaan RA, Madden CJ, Yaxley RH and Aveyard ME (2004) Moving words: Dynamic representations in language comprehension. *Cognitive Science* **28**, 611–619.

Zwaan RA and Pecher D (2012) Revisiting mental simulation in language comprehension: Six replication attempts. *PloS one* **7**, e51382.

Zwaan RA, Stanfield RA and Yaxley RH (2002) Language comprehenders mentally represent the shapes of objects. *Psychological Sience* **13**, 168–171.

https://doi.org/10.1017/S1366728921001115 Published online by Cambridge University Press