Journal of Psycholinguistic Research
https://doi.org/10.1007/s10936-019-09670-x

Perceptual Representations in L1, L2 and L3 Comprehension: Delayed Sentence–Picture Verification

Donggui Chen[1,2] · Ruiming Wang[1] · Jinqiao Zhang[3] · Cong Liu[4]

© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract
We used delayed sentence–picture verification tasks to investigate multilingual perceptual representations. In experiment 1, participants listened to sentences with implied shapes. After a 10-min interval, they judged whether pictures had been mentioned in the preceding sentences or not. Results in experiment 1 showed significant match effect in L1, but not in high proficient L2 or low proficient L3. In experiment 2, Participants listened to one language block, then immediately judged one picture block, totally three language–picture blocks. Results in experiment 2 were parallel to results in experiment 1. Our study supports the view of distributed conception: L2 and L3 are associated with less perceptual symbols than L1, indicating great impact of acquisition styles on perceptual representations. Our results show little impact of language proficiency levels on perceptual representations in delayed tasks.

Keywords Perceptual representations · Multilingual · Embodiment · Language comprehension

Introduction

Frequent international contact has increased multilingual populations worldwide. People learn their first language (L1) very effectively. Routinely every normal child can master his or her L1 within few years after birth. But learning a second (L2) or a foreign language (L3) is much more struggling. In many cases, L2 and L3 learning may end up with a faulty and inexact mastery.

✉ Ruiming Wang
wangrm@scnu.edu.cn

[1] Guangdong Provincial Key Laboratory of Mental Health and Cognitive Science, Center for Studies of Psychological Application, School of Psychology, South China Normal University, Zhongshan Street West, 55, Tianhe District, Guangzhou 510613, China

[2] School of Foreign Languages, Guangdong Polytechnic Normal University, Guangzhou, China

[3] College of Chinese Language and Culture, Jinan University, Guangzhou, China

[4] Department of Psychology, Qingdao University, Qingdao 266071, China

Why can L1 learning be so much more effective than L2 or L3 learning? In the theory of embodiment, linguistic symbols are not arbitrarily related to their outside world referents. Instead, linguistic symbols are embodied by perceptual experiences (Barsalou 2008). On this view, language learning is not only the acquisition of vocabulary, grammar or semantic system, but also the integration of perceptual experiences and the linguistic meaning.

L1 is acquired through years of everyday interaction with people in real and various situations. L1 acquisition is typically combined with specific gestures, eye-movements and physical orientations towards the described entity in real time communication (Glenberg and Gallese 2012). L1 acquisition is a sensorimotor process. For children, L1 acquisition is not only the connection of words and referents, or the building up of complex grammar system. More importantly, L1 acquisition is accompanied with children's cognitive development. In short, the daily-based acquisition style promotes the establishment of strong perception-conception links in L1 comprehension. It is reasonable to believe that people represent L1 with strong perceptual symbols.

However, learning L2 or L3 in classroom setting is different. L2 or L3 is usually taught to elder children, teenagers or even adults, when their L1 have already dominated in mind. School learning takes place in specific and limited settings. School learning provides less interactions with language and physical experiences during language learning phase. People typically learn a foreign word form with the help of its translation equivalent. Therefore, perceptual representations in L2 and L3 could be much weaker compared with those in L1.

The role of perceptual experiences and their neural systems in the construction of linguistic meaning had received a great deal of attention during the past two decades. Many studies have provided converging evidence to support perceptual representations in L1 comprehension. Evidence for L1 embodiment was obtained by many behavioral studies. Visual simulations of object properties (e.g., shape, orientation, etc.) were found during sentence comprehension (e.g., Stanfield and Zwaan 2001; Zwaan et al. 2002). Motor system facilitated language understanding and memory (e.g., Fischer and Zwaan 2008; Glenberg and Gallese 2012; Marino et al. 2013). Emotional words evoked a series of bodily reactions (e.g., Davis et al. 2015; Fino et al. 2016; Baumeister et al. 2017). Recently, growing clinical and neuropsychological studies lent support to L1 embodiment (e.g., Hauk et al. 2004; Fernandino et al. 2013; Buccino et al. 2017; Vukovic et al. 2017; Birba et al. 2017; Cotelli et al. 2018; Gijssels et al. 2018). There were also literature review and meta-analysis for L1 embodiment (e.g., Pulvermüller 2013; Buccino et al. 2016). The previous studies were mainly about L1 perceptual representations during immediate language processing. The main disagreement lied in whether L1 was strongly embodied or moderately embodied (Horchak et al. 2014; Santos et al. 2011). Thus, more studies are needed to explore perceptual representations in L1 during delayed language processing to find answer to this disagreement.

Whether L2 and L3 are perceptually represented still remains an open question. In theory, different models have been proposed to account for bilingual language representations. In empirical studies, controversial results have been found in investigating L2 perceptual representations.

The Concept Mediation Model (Potter et al. 1984) assumed that L2 lexical had direct access to the conceptual store, without the help of L1 lexicon. The Bilingual Interactive Activation plus Model (Dijkstra and Van Heuven 2002) assumed that both target and non-target languages could be automatically activated even in a pure monolingual mode. L1 and L2 interacted with each other on the conceptual level. Based on these two models, L2 could have a direct connection to the experiential representations. This is a view of shared conception. It assumes that the perceptual representations in L1 could transfer equally into L2 processing through the shared conception. This assumption was supported by some behavioral and even neuropsychological studies (e.g., Bergen et al. 2010; De Grauwe et al. 2014; Dudschig et al. 2014; Repetto et al. 2017; Buccino et al. 2017; García-Gámez and Macizo 2018; Bechtold et al. 2018).

For example, in a behavioral experiment, researchers used stroop task to investigate the activation of sensorimotor simulation in L2 comprehension. Participants saw L1 or L2 words referring to entities with a typical location (e.g., star, mole) or to an emotion (e.g., happy, sad). Participants responded to the words' ink color with an upward or downward arm movement. Researchers found that motor responses were automatically activated in L2 processing similar to that in L1 processing (Dudschig et al. 2014). Also, in an fMRI experiment, researchers investigated whether non-native semantic representations were rich enough to allow for activation in motor and somatosensory brain areas. Participants made lexical decisions about visually presented motor and non-motor verbs. Region-of-interest (ROI) and whole-brain analyses indicated that L2 speakers, similar to L1 speakers, showed significantly increased activation for simple motor compared to non-motor verbs in motor and somatosensory regions (De Grauwe et al. 2014). These studies showed that there were automatically activated perceptual responses in L2 processing similar to L1 processing.

However, the Word Association Model (Scarborough et al. 1984) assumed that L2 comprehension had always to be mediated by L1 lexical entries to access conceptual store. The Revised Hierarchical Model (Kroll and Stewart 1994) assumed that language proficiency played an important role. Early L2 learners needed L1 mediation, while L2 learners with high proficiency could access the conceptual store without reliance on L1 lexicon. Based on these two models, L2 could have an indirect connection to the perceptual representations. Language proficiency, age of acquisition (AoA) and L2 exposure could play important roles in L2 perceptual representations.

While, the Distributed Conceptual Feature Model (De Groot 1992) supposed that some words (e.g., father) had exactly the same meaning in L1 and L2, while other words (e.g., beauty) differed somewhat in the two languages. Likewise, the Sense Model (Finkbeiner et al. 2004) supposed that words in L1 were associated with a greater number of semantic senses than words in L2. L2 representations were less developed than L1 representations. There should be just partially overlapping semantic representations for L1 and L2 words.

Taken together, these four bilingual representation models provide a view of distributed conception. It assumed that perceptual representations in L2 were less rich than those in L1 (Finkbeiner et al. 2004; Monaco et al. 2019). Thus, the behavioral observation and neural activation for L2 perceptual representations must be reduced in comparison with L1, or even be absent. This assumption was also supported by growing behavioral and neuropsychological studies (e.g., Vukovic and Shtyrov 2014; Hsu et al. 2015; Caldwell-Harris 2015; Foroni 2015; Xue et al. 2015; Sheikh and Titone 2015; Qian 2016; Ahlberg et al. 2017; Pavlenko 2017; Baumeister et al. 2017).

For example, in an eye-tracking experiment, bilingual participants naturally read sentences (with positive/neutral/negative words) in their L2 and then were compared with L1 readers. Their results showed that emotional advantage was only found in positive words, suggesting only positive words were emotionally embodied, but not negative words (Sheikh and Titone 2015). Also, in a facial measure experiment, researchers found that the zygomatic muscle was activated when reading affirmative sentences (e.g., I am smiling) relevant to the muscle. In contrast, reading sentences in the negative form (e.g., I am not smiling) did not lead to relaxation or inhibition of the zygomatic muscle (Foroni 2015). In a neurophysiological study as well, researchers used high-density EEG to dynamically measure changes in the cortical motor system's activity, indexed by event-related desynchronization (ERD) of the mu-rhythm, in response to passively reading L1 and L2 action words. They found that mu-rhythm ERD was significantly stronger for L1 than L2 words. Their results revealed important differences in L2 sensorimotor brain activity (Vukovic and Shtyrov 2014). These studies showed that there were quantitative differences between L1 and L2 perceptual representations during language comprehension. L2 is less embodied than L1 or as a different pattern.

Language proficiency, age of acquisition, and L2 exposure were the main factors to modulate the degree of L2 embodiment, according to the previous theories and studies (e.g. Kroll and Stewart 1994; Qian 2016; Ahlberg et al. 2017; Monaco et al. 2019). Some recent studies also showed that different types of L2 learning (e.g., learning with gestures, pictures or manipulation) had different learning effect on memory performances and activation of motor areas in the brain (e.g., Repetto et al. 2017; Bechtold et al. 2018; García-Gámez and Macizo 2018). More studies are needed to investigate the degree of L2 and L3 embodiment, and explore the factors influencing the degree of L2 embodiment.

The Present Study

Previous researches for perceptual representations were mainly about immediate language processing. The present study used the delayed sentence–picture verification paradigm (Pecher et al. 2009) to investigate perceptual representations on multilingual participants. The present study aimed at investigating whether L1 was strongly and lastingly embodied, and whether high proficient L2 could be embodied as strongly as L1. It also investigated whether low proficient L3 could be somehow embodied. If perceptual representations were found across languages regardless of language proficiency, then the shared conception view was supported. If the finding of perceptual representations could be mediated by language proficiency or acquisition styles, the distributed conception view was supported. We predicted that L1 and high proficient L2 were perceptually represented, and L1 had stronger embodiment than L2, and low proficient L3 had no embodiment.

The delayed sentence–picture verification paradigm (Pecher et al. 2009) had been used to explore how fine-grained perceptual features were activated during delayed language comprehension. In this paradigm, participants were asked to verify whether objects in the pictures were mentioned in a preceding sentence or not. The key feature of the paradigm was that the sentenced object could either match or mismatch the same pictured object in its perceptual features. For example, the implied shape of an eagle in a sentence (There is an eagle in the sky) matched or mismatched the shape of an eagle in a picture (eagle with wings stretched out or with wings drawn in). Critically, the target perceptual features were task-irrelevant and only implied in the sentences. If participants performed perceptual simulations, they would show sensitivity to the implied features. In other words, they were faster to verify pictures that matched the perceptual features than that mismatched. Reaction times were typically facilitated when the implied features matched the visually presented targets (i.e., the match-effect: Stanfield and Zwaan 2001).

Participants are native Cantonese speakers recruited from South China Normal University in Guangzhou. They were all born, grew up, and are still living and studying in Cantonese dialect area. Cantonese is the mother tongue and the dominate language used in everyday life. In our study, Cantonese was considered as L1 with daily-based acquisition. How to determine the language proficiency of L2 is a questionable issue in previous studies. Our participants began to learn Mandarin formally as official language at about

7 years old when going to primary schools. As Chinese college students, our participants have been speaking and using Mandarin in formal settings for more than 10 years, indicating very high L2 proficiency level. To make it more scientific, we asked the participants to do a 7-point-self-rating scale of language proficiency for Mandarin. The mean Mandarin proficiency is $5.94 \pm .75$. Thus, Mandarin was considered as the high proficient L2 in the present study. In addition, our participants speak English as foreign language mainly in classroom and other occasional settings. They are non-English majors, and have passed the College English Test-Band 4 (a national test for Non-English majors in China), but not yet passed the College English Test-Band 6, showing their relatively lower English proficiency level. The mean English proficiency of 7-point-self-rating scale is $3.44 \pm .65$. Thus, English was considered as the low proficient L3 in our study. The difference between their L2 and L3 proficiency was significant: $t = 13.533$, $p = .001$.

The rationale of our research goes as follows. If L1, L2 and L3 share the same conception system, then perceptual representations could be found equally in these three languages. That's L1 = L2 = L3. If L1, L2 and L3 do not or only partially share conceptual system, then we should expect decreasing perceptual representations in L1, L2 and L3. That's L1 > L2 > L3, when language proficiency levels can mediate the access of linguistic entries to perception. And we should expect that embodiment in daily-based L1 will be stronger than that in school-based L2 and L3. That's L1 > L2 = L3, when acquisition styles determine the access of linguistic entries to perception system.

Experiment 1

Participants

The participants were 36 native Cantonese speakers (23 females, 13 males, aged 18–22) from South China Normal University in Guangzhou. They speak Mandarin as official language, and English as foreign language. They are non-English majors and have passed the College English Test-Band 4 but not College English Test-Band 6. They all reported have normal hearing ability and no reading or imaging impairment. They were paid 35RMB to take part in the study.

Materials

We used 60 experimental sets. Each experimental set consisted of 6 meaningful sentences and 2 black and white pictures (see Table 1). The 6 sentences described one object with 2 different implied shapes, in 3 language versions (Cantonese/Mandarin/English). We took 24 experimental materials from the previous study (Zwaan et al. 2002) and added 36 similar materials.

Also, we used 120 filler sentences (30 meaningful and 90 meaningless sentences, in 3 language versions) and 60 unrelated filler pictures. Both experimental and filler sentences were recorded into audio version by native Cantonese, native Mandarin and native English speakers respectively.

12 lists were created, each list with 180 sentences divided into 3 language blocks. Each language block consisted of 20 meaningful experimental sentences, 10 meaningful filler sentences, and 30 meaningless filler sentences. Also, each list consisted of 120 pictures, including 60 related pictures and 60 unrelated pictures. Each participant received only

Table 1 Material examples for each experimental set

Languages	Audio Sentences	Pictures
L1:Cantonese	个天度飞住只鹰。	
	鹰窦入边有只鹰。	
L2:Mandarin	天空中有只老鹰。	
	巢穴中有只老鹰。	
L3:English	There is an eagle in the sky.	
	There is an eagle in the nest.	

Cantonese has no written form. The transliterated sentences in the table are just for illustration. All the sentence materials are in audio version

one sentence and one picture from each experimental set. Meaningful sentences required a "Yes" response, and meaningless sentences required a "No" response. Related pictures required a "Yes" response, and unrelated pictures required a "No" response. In addition, 6 sentences (3 meaningful and 3 meaningless) and 6 pictures (3 related and 3 unrelated) were used in practice trails.

Before the experiment, 15 participants who did not participate in the experiments rated the comprehensibility and imageability of all materials (pictures and sentences) on the likert 5-point scale. Mean score for comprehensibility is 4.89 and mean score for imageability is 4.14.

Procedure

The entire experiment was displayed in a computer screen using Eprime. The delayed picture verification paradigm (Pecher et al. 2009) consisted of three phases (see Fig. 1). Firstly,

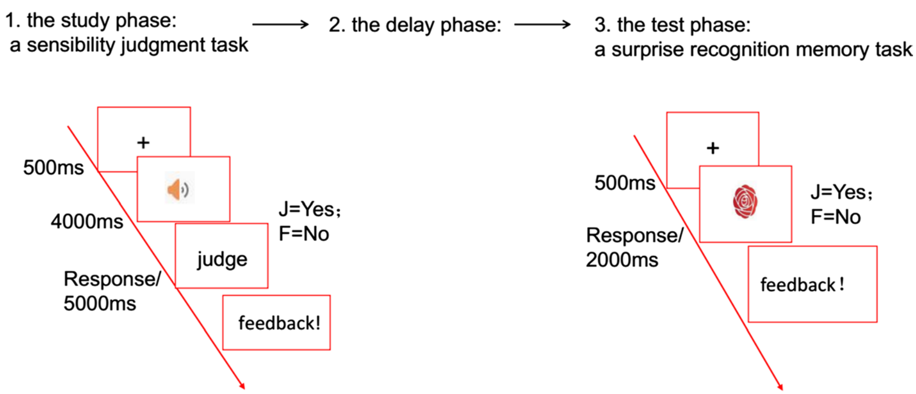

Fig. 1 Procedure of experiment 1

in the study phase participants listened and decided whether each sentence was meaningful or not. 60 sentences were presented in random order within each language block, and 3 language blocks were counterbalanced. Secondly, in the delay phase participants were asked to fill in the forms about their language background. The delay phase lasted for about 10 m. Thirdly, in the test phase participants saw and judged whether the object in a picture had been mentioned in any sentence that they had heard in the study phase. The 120 pictures were presented in random order. Experiment 1 lasted for about 45 m.

Results and Discussion

Fillers and practice items were removed from the data. Mean hit rates of picture recognition were calculated for each participant. Data from 2 participants were deleted, because their mean hit rates was below 60%. The reaction time of incorrect responses (25.6%) were excluded. For each condition (3 language types × 2 match types), trials slower than mean reaction time plus 3 × standard deviation (1.9%) were also excluded from the analysis. Then, hit rates and reaction times were analyzed separately.

The mean hit rates for 6 conditions in the picture recognition task were presented in Table 2. A 3 (L1: Cantonese vs. L2: Mandarin vs. L3: English) × 2 (match vs. mismatch) repeated measures ANOVA was performed on the hit rates. There was a significant effect between language types, $F(2,66) = 12.048$, $p = .000$. There was no significant effect between match types, $F(1,33) = .129$, $p = .722$. There was no interaction effect between language types and match types, $F(2,66) = 1.722$, $p = .187$. Analysis of the errors showed no speed-for-accuracy trade off.

For different conditions, the mean reaction times in the picture recognition task were presented in Table 3 and Fig. 2.

We reported the pattern of results based on mean RTs after excluding trials beyond 3 standard deviation, since this is a common practice in previous studies. A 3 (L1: Cantonese vs. L2: Mandarin vs. L3: English) × 2 (match vs. mismatch) repeated measures ANOVA was performed on the mean RTs. There was a significant effect between language types, $F(2,66) = 3.719$, $p = .029$. More importantly, there was a significant effect between match types, $F(1,33) = 5.894$, $p = .021$. Further simple effects analyses found that the match effect was significant within L1, but not within L2 or L3 (see Table 4). There was no interaction effect between language types and match types, $F(2,66) = 1.795$, $p = .174$.

The aim of experiment 1 is to investigate whether and how perceptual representations exist across L1, L2 and L3 comprehension. More specifically, whether L1 perceptual representations are strongly enough to be found in long term memory state; whether high proficient L2 is perceptually represented; whether low proficient L3 is perceptually represented.

As predicted, results in experiment 1 showed a significant main effect in match types, indicating the existence of perceptual representations. Also, it showed a significant main

Table 2 Mean hit rates in the picture recognition task as a function of match and language type in experiment 1

	L1: Cantonese		L2: Mandarin		L3: English	
	Match	Mismatch	Match	Mismatch	Match	Mismatch
Hit rates (SD)	.788 ± .155	.771 ± .119	.785 ± .133	.747 ± .131	.668 ± .157	.703 ± .145

Table 3 Mean reaction times in the picture recognition task as a function of match and language type in experiment 1 (ms)

	L1: Cantonese		L2: Mandarin		L3: English	
	Match	Mismatch	Match	Mismatch	Match	Mismatch
RT (SD)	1029.28 ± 239.57	1120.41 ± 240.82	1128.63 ± 265.08	1139.46 ± 229.61	1090.65 ± 250.90	1119.21 ± 239.28

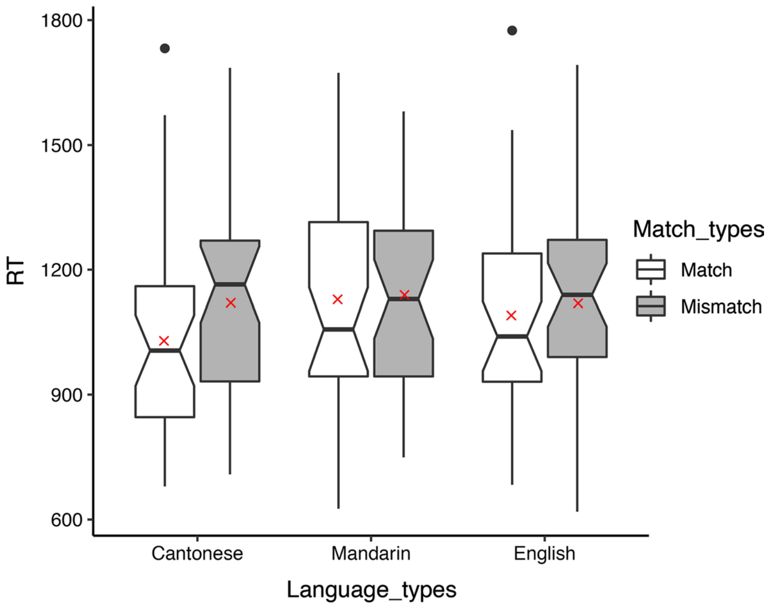

Fig. 2 Reaction times in the picture recognition task in experiment 1

Table 4 Test involving match effect within language types in experiment 1

Languages	SS	DF	MS	F	Sig of F
Match effect within L1: Cantonese	141,195.11	1	141,195.11	11.04	.002**
Match effect within L2: Mandarin	1994.45	1	1994.45	.1	.759
Match effect within L3: English	13,866.88	1	13,866.88	.84	.365

effect in language types, indicating perceptual representations existed differently across these languages.

When we took a further step to analyze the match effect within language types, we found that match effect only existed during L1 comprehension, but not in high proficient L2 or low proficient L3. That was L1 > L2 = L3 = 0.

The results showed a strong match effect in L1, even though the experiment involved sentence processing and delayed picture recognition. The finding of match effect in delayed L1 comprehension was in accordance with earlier studies, showing the existence of perceptual representations in L1 (e.g., Stanfield and Zwaan 2001; Zwaan 2016; Baumeister et al. 2017; Gijssels et al. 2018; Cotelli et al. 2018). Our finding further confirmed that L1 comprehension was strongly and lastingly embodied.

Unexpectedly, results in experiment 1 showed no match effect during high proficient L2 comprehension. The absent of match effect suggested no perceptual representations during delayed L2 comprehension. This result did not support the view of shared conception which suggested perceptual representations in L1 processing could transfer equally to L2 processing. Rather, this result lent support to the view of distributed conception which suggested L2 perceptual representations could be much weaker or even were absent. In addition, no match effect was found during low proficient L3 comprehension, indicating no perceptual representations in delayed L3 comprehension. Since no match effect was found in high proficient L2 processing, this result suggested that language proficiency was not an impact factor for perceptual representations during delayed language processing.

In short, the results of experiment 1 showed that L1 had strong perceptual representations, while L2 and L3 had no perceptual presentations during delayed language processing. The results indicated that perceptual representations were mainly affected by acquisition styles, because perceptual representations were only found in daily-based L1, but not in school-based L2 or L3. Further, the results indicated that L1, L2 or L3 did not share conception system.

It is reasonable to find strong perceptual representations in L1. It is also reasonable to find no perceptual representations in low proficient L3. However, it is surprising to find no perceptual representations in high proficient L2 comprehension. Especially, in the present study, our participants are college students from Mainland China. Their high proficiency in Mandarin is indisputable. Why no perceptual representations were found during delayed Mandarin comprehension? We considered there might be other potential influences on the display of perceptual representations during delayed language comprehension. The match effect might be modulated by the duration and complexity of experimental tasks. Therefore, we run experiment 2 to examine whether the match effect would be observed in L2 and L3 in a shorter and easier experiment.

Experiment 2

Participants

Another 36 participants (29 females, 7 males, aged 18–22) were recruited under the same standard as that in experiment 1.

Materials

Materials were the same as those in Experiment 1.

Procedure

In experiment 2, we deleted the delay phase, and divided the pictures into 3 blocks and moved the picture blocks up to follow the relevant language blocks (see Fig. 3). In this way, participants first listened to and judged 60 sentences in each language block. Then immediately after a language block, they saw and judged 40 pictures. There were 3 language blocks followed by 3 picture blocks respectively. Experiment 2 lasted for about 35 m.

Results and Discussion

Fillers and practice items were removed from the data. Data from 1 participant were deleted, because the mean hit rate was below 60%. The reaction time of incorrect responses (19%) were excluded. For each condition, trials slower than mean reaction time plus

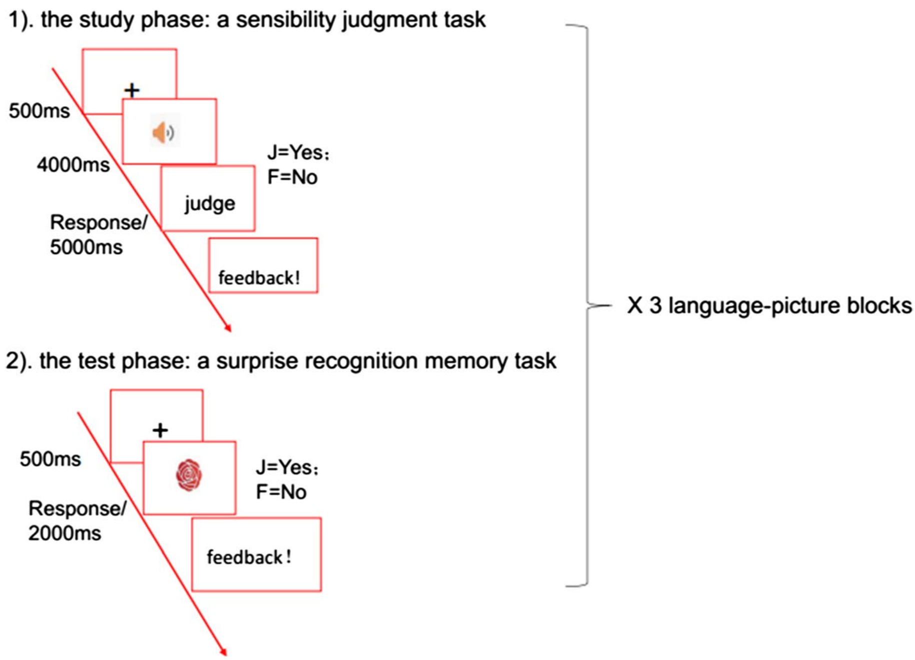

Fig. 3 Procedure of experiment 2

3 × standard deviation (1.7%) were also excluded from the analysis. For different conditions, the mean hit rates in the picture recognition task were presented in Table 5.

A 3 (L1: Cantonese vs. L2: Mandarin vs. L3: English) × 2 (match vs. mismatch) repeated measures ANOVA was performed on the hit rates. There was a significant effect between language types, F (2,68) = 9.126, $p < .001$. There was no significant effect between match types, F (1,35) = .223, $p = .276$. There was no interaction effect between language types and match types, F (2,68) = .131, $p = .878$.

For different conditions, the mean reaction times in the picture recognition task were presented in Table 6 and Fig. 4.

A 3 (L1: Cantonese vs. L2: Mandarin vs. L3: English) × 2 (match vs. mismatch) repeated measures ANOVA was performed on the reaction times. There was a significant effect between language types, F (2,68) = 7.470, $p = .001$. More importantly, there was a significant effect between match types, F (1,34) = 5.719, $p = .022$. There was a significant interaction effect between language types and match types, F (2,68) = 4.198, $p = .019$. Further simple effects analyses found that the match effect was significant only within L1, but not within L2 or L3 (see Table 7). In short, the results of experiment 2 further confirmed the results of experiment 1.

Table 5 Mean hit rates in the picture recognition task as a function of match and language type in experiment 2

	L1: Cantonese		L2: Mandarin		L3: English	
	Match	Mismatch	Match	Mismatch	Match	Mismatch
Hit rates (SD)	.846 ± .174	.863 ± .144	.809 ± .160	.837 ± .142	.746 ± .167	.754 ± .174

Table 6 Mean reaction times in the picture recognition task as a function of match and language type in experiment 2 (ms)

	L1: Cantonese		L2: Mandarin		L3: English	
	Match	Mismatch	Match	Mismatch	Match	Mismatch
RT (SD)	999.14 ± 210.27	1131.20 ± 254.46	1102.52 ± 232.97	1120.66 ± 214.10	1185.21 ± 300.65	1210.20 ± 269.28

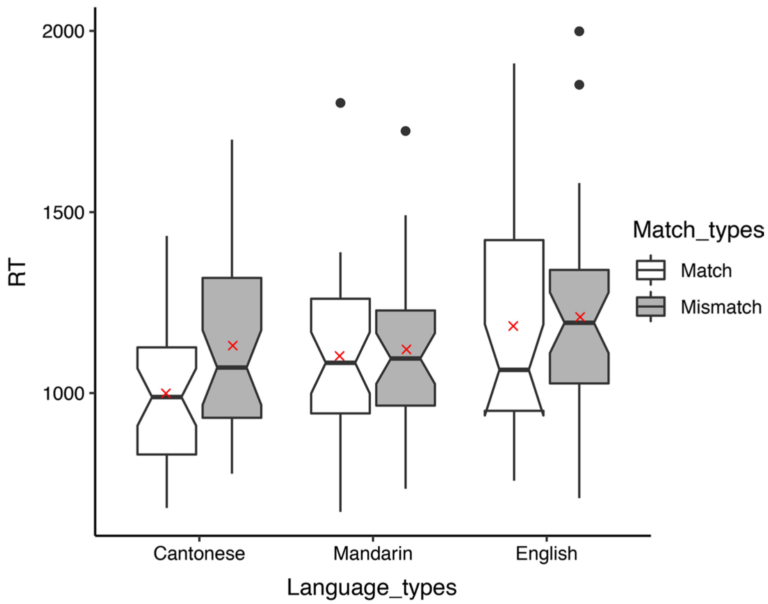

Fig. 4 Reaction times in the picture recognition task in experiment 2

The aim of experiment 2 is to investigate whether the match effect can be observed in high proficient L2 comprehension in a shorter and easier experiment. Results in experiment 2 showed a significant main effect in match types, indicating the existence of perceptual representations. Also, it showed a significant main effect in language types, indicating perceptual representations existed differently across these languages.

We took a further step to analyze the match effect within language types. Parallel to experiment 1, we found perceptual representations only in L1, but not in L2 or L3 comprehension, even though we had reduced the duration and complexity in experiment 2. Again, that was $L1 > L2 = L3 = 0$. Experiment 2 further confirmed our findings in experiment 1, showing that, in delayed language processing, perceptual representations existed only in daily-based L1, but not in school-based L2 or L3.

General Discussion

The purpose of the present study was to investigate whether and how perceptual representations were involved in delayed language comprehension. We conducted two experiments using delayed sentence–picture verification paradigm on multilingual participants.

Table 7 Test involving match effect within language type in experiment 2

Languages	SS	DF	MS	F	Sig of F
Match effect within L1: Cantonese	305,173.84	1	305,173.84	16.77	.000***
Match effect within L2: Mandarin	5759.98	1	5759.98	.28	.599
Match effect within L3: English	10,930.71	1	10,930.71	.41	.526

In experiment 1, participants listened to and judged 180 sentences in three language blocks and then filled the questionnaires, and finally saw and judged 120 pictures. Their recognition for pictures was significantly better if the picture matched the implied shape of the object in an earlier L1 sentence. But no match effect was found in either L2 or L3 comprehension. In experiment 2, participants listened to and judged 60 sentences in each language block, and then immediately saw 40 pictures. Totally there were 3 language blocks followed by 3 picture blocks respectively. Again, significant match effect was found only in L1, but not in L2 or L3 comprehension.

Taken together, results in experiment 1 and 2 showed that, in delayed language processing, perceptual representations were only observed in daily-based L1, but not in school-based L2 or L3. Specifically, perceptual representations were not observed in either high proficient L2 or low proficient L3. Our study indicated that L1, L2 and L3 did not share the same conception system. Also, our study indicated no impact of language proficiency levels on perceptual representations during delayed language processing.

Our finding of perceptual representations in L1 comprehension was similar to that in recent studies (Pecher et al. 2009; Nijhof and Willems 2015; Zwaan 2016), further confirming that L1 was strongly embodied.

However, we found no embodiment in high proficient L2 processing, which was not in line with the previous studies. We consider that previous studies mainly focus on immediate L2 word processing, while we focus on delayed semantic processing. Moreover, previous studies mostly involved some specific word types, for example homophones (Vukovic and Williams 2014), cognate words (De Grauwe et al. 2014), action words (García-Gámez and Macizo 2018) or emotional words (Foroni 2015). These specific word types could be easier to trigger embodied reactions.

We considered that there were two possibilities for the absence of perceptual representations in delayed L2 and L3 comprehension. One possible explanation is that perceptual representations do not exist at all in non-L1 language comprehension. Another possible explanation is that the display of non-L1 perceptual representations is under certain constraints.

We tend to support the second possibility. Firstly, participants in our study are still living and studying in Cantonese dialect area. The dominance of L1 may suppress the display of perceptual representations in L2, which is mainly used in formal settings, let alone L3, which is just occasionally used.

Secondly, our study focused on perceptual representations in the long-term memory state. Memory state may constrain the display of non-L1 perceptual representations.

Thirdly, the complexity of the tasks may suppress the display of non-L1 perceptual representations. Participants did lots of switch works in our study: switch among multiple languages, switch from auditory modality (listening to sentences) to visual modality (seeing pictures), and switch from sensibility judgment task to picture recognition task.

Finally, behavioral studies are not sensitive enough to completely refuse the existence of non-L1 perceptual representations (De Grauwe et al. 2014; Vukovic and Shtyrov 2014). Neuropsychological studies are needed in the near future to explore the non-L1 perceptual representations.

To conclude, our findings demonstrated that delayed L1, L2 and L3 comprehension were represented differently on multilingual individuals. Our results lent support to the view of distributed conception: L2 and L3 were associated with less semantic senses and perceptual symbols than L1. Perceptual representations in L2 and L3 were much weaker compared to that in L1 or even were absent. Our results indicated great impact of acquisition styles on perceptual representations. We argued that the display of non-L1 perceptual representations was constrained under certain conditions.

We found that language proficiency levels had no effect on perceptual representations during long-term memory state. Future studies are needed to further investigate whether language proficiency levels can impact the display of perceptual representations in working memory state. Our results were picked up via auditory and visual stimuli, more studies are needed to explore the response on olfactory or gustatory stimuli. Also, a recent paper stressed differential coding of perception in the world's languages (Majid et al. 2018). More studies are needed to investigate the issue of language relativity.

Funding This study was funded by the National Social Science Foundation of China (19AYY009).

Compliance with Ethical Standards

Conflict of interest We declare that we have no conflict of interest.

References

Ahlberg, D. K., Bischo, H., Kaup, B., Bryant, D., & Strozyk, J. V. (2017). Grounded cognition: Comparing language × space interactions in first language and second language. *Applied Psycholinguistics, 39,* 437–459.

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

Baumeister, J. C., Foroni, F., Conrad, M., Rumiati, R. I., & Winkielman, P. (2017). Embodiment and emotional memory in first vs. second language. *Frontiers in Psychology, 8,* 394.

Bechtold, L., Ghio, M., Lange, J., & Bellebaum, C. (2018). Event-related desynchronization of mu and beta oscillations during the processing of novel tool names. *Brain and Language, 177–178,* 44–55.

Bergen, B., Lau, T.-T. C., Narayan, S., Stojanovic, D., & Wheeler, K. (2010). Body part representations in verbal semantics. *Memory & Cognition, 38,* 969–981.

Birba, A., García-Cordero, I., Kozono, G., Legaz, A., Ibáñez, A., Sedeño, L., et al. (2017). Losing ground: Frontostriatal atrophy disrupts language embodiment in Parkinson's and Huntington's disease. *Neuroscience and Biobehavioral Reviews, 80,* 673–687.

Buccino, G., Colagè, I., Gobbi, N., & Bonaccorso, G. (2016). Grounding meaning in experience: A broad perspective on embodied language. *Neuroscience and Biobehavioral Review, 69,* 69–78.

Buccino, G., Dalla Volta, R., Arabia, G., Morelli, M., Chiriaco, C., Lupo, A., et al. (2017). Processing graspable object images and their nouns is impaired in Parkinson's disease patients. *Cortex, 100,* 32–39.

Caldwell-Harris, C. L. (2015). Emotionality differences between a native and foreign language: Implications for everyday life. *Current Directions in Psychological Science, 24,* 214–219.

Cotelli, M., Manenti, R., Brambilla, M., & Borroni, B. (2018). The role of the motor system in action naming in patients with neurodegenerative extra-pyramidal syndromes. *Cortex, 100,* 191–214.

Davis, J. D., Winkielman, P., & Coulson, S. (2015). Facial action and emotional language: ERP evidence that blocking facial feedback selectively impairs sentence comprehension. *Journal of Cognitive Neuroscience, 27,* 2269–2280.

De Grauwe, S., Willems, R. M., Rueschemeyer, S. A., Lemhöfer, K., & Schriefers, E. (2014). Embodied language in first- and second-language speakers: Neural correlates of processing motor verbs. *Neuropsychologia, 56,* 334–349.

De Groot, A. M. B. (1992). Determinants of word translation. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 18,* 1001–1018.

Dijkstra, T., & Van Heuven, W. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5,* 175–197.

Dudschig, C., de la Vega, I., & 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.

Fernandino, L., Conant, L. L., Binder, J. R., Blindauer, K., Hiner, B., Spangler, K., et al. (2013). Where is the action? Action sentence processing in Parkinson's disease. *Neuropsychologia, 51,* 1510–1517.

Finkbeiner, M., Forster, K., Nicol, J., & Nakamura, K. (2004). The role of polysemy in masked semantic and translation priming. *Journal of Memory and Language, 51,* 1–22.

Fino, E., Menegatti, M., Avenanti, A., & Rubini, M. (2016). Enjoying vs. smiling: Facial muscular activation in response to emotional language. *Biological Psychology, 118,* 126–135.

Fischer, M. H., & Zwaan, R. A. (2008). Grounding cognition in perception and action. *Quarterly Journal of Experimental Psychology, 61,* 825–857.

Foroni, F. (2015). Do we embody second language? Evidence for 'partial' simulation during processing of a second language. *Brain and Cognition, 99,* 8–16.

García-Gámez, A. B., & Macizo, P. (2018). Learning nouns and verbs in a foreign language: The role of gestures. *Applied Psycholinguistics, 40,* 473–507.

Gijssels, T., Ivry, R. B., & Casasanto, D. (2018). tDCS to premotor cortex changes action verb understanding: Complementary effects of inhibitory and excitatory stimulation. *Scientific Reports, 8,* 11452.

Glenberg, A. M., & Gallese, V. (2012). Action-based language: A theory of language acquisition, comprehension, and production. *Cortex, 48,* 905–922.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatopic representation of action words in human motor and premotor cortex. *Neuron, 41,* 301–307.

Horchak, O. V., Giger, J. C., Cabral, M., & Pochwatko, G. (2014). From demonstration to theory in embodied language comprehension: A review. *Cognitive Systems Research, 29–30,* 66–85.

Hsu, C. T., Jacobs, A. M., & 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.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33,* 149–174.

Majid, A., Roberts, S. G., Cilissen, L., et al. (2018). Differential coding of perception in the world's languages. *Proceedings of the National Academy of Sciences, 115,* 11369–11376.

Marino, B. F., Gough, P. M., Gallese, V., Riggio, L., & Buccino, G. (2013). How the motor system handles nouns: A behavioral study. *Psychological Research, 77,* 64–73.

Monaco, E., Jost, L. B., Gygax, P. M., & Annoni, J. M. (2019). Embodied semantics in a second language: Critical review and clinical implications. *Frontiers in Human Neuroscience, 13,* 110.

Nijhof, A. D., & Willems, R. M. (2015). Simulating fiction: Individual differences in literature comprehension revealed with fMRI. *PLoS ONE, 10,* 1–17.

Pavlenko, A. (2017). Do you wish to waive your rights? Affect and decision-making in multilingual speakers. *Current Opinion in Psychology, 17,* 74–78.

Pecher, D., Van Dantzig, S., Zwaan, R. A., & Zeelenberg, R. (2009). Language comprehenders retain implied shape and orientation of objects. *Quarterly Journal of Experimental Psychology, 62,* 1108–1114.

Potter, M. C., So, K.-F., Von Eckardt, B., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior, 23,* 23–38.

Pulvermüller, F. (2013). Semantic embodiment, disembodiment or misembodiment? In search of meaning in modules and neuron circuits. *Brain and Language, 127,* 86–103.

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,* 484–494.

Repetto, C., Pedroli, E., & Macedonia, M. (2017). Enrichment effects of gestures and pictures on abstract words in a second language. *Frontiers of Psychology, 8,* 2136.

Santos, A., Chaigneau, S. E., Simmons, W. K., & Barsalou, L. W. (2011). Property generation reflects word association and situated simulation. *Language and Cognition, 3,* 83–119.

Scarborough, D. L., Gerard, L., & Cortese, C. (1984). Independence of lexical access in bilingual word recognition. *Journal of Verbal Learning and Verbal Behavior, 23,* 84–99.

Sheikh, N., & Titone, D. (2015). The embodiment of emotional words in a second language: An eye-movement study. *Cognition and Emotion, 30,* 488–500.

Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science, 12,* 153–156.

Vukovic, N., Feurra, M., Shpektor, A., Myachykov, A., & Shtyrov, Y. (2017). Primary motor cortex functionally contributes to language comprehension: An online rTMS study. *Neuropsychologia, 96,* 222–229.

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

Vukovic, N., & Williams, J. N. (2014). Automatic perceptual simulation of first language meanings during second language sentence processing in bilinguals. *Acta Psychologica, 145,* 98–103.

Xue, J., Marmolejo-Ramos, F., & Pei, X. (2015). The linguistic context effects on the processing of body-object interaction words: An ERP study on second language learners. *Brain Research, 1613,* 37–48.

Zwaan, R. A. (2016). Situation models, mental simulations, and abstract concepts in discourse comprehension. *Psychonomic Bulletin & Review, 23,* 1028–1034.

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

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.