Article

Listening for Imagery by Native Speakers and L2 Learners

Shu-Ling Wu

Department of Languages, Cultures, and International Trade, Southern Illinois University Carbondale, 1000 Faner Drive, Carbondale, IL 62901, USA; shulingwu@siu.edu; Tel.: +1-618-453-5411

Abstract: Slobin's thinking-for-speaking (TFS) hypothesis suggests that speakers are habitually attuned to aspects of an event that are readily codable in the language while they are formulating speech. This TFS process varies considerably cross-linguistically and can be observed in all forms of production and reception including listening for understanding or mental imagery. This study explored whether second language learners (L2) engage in mental simulation of deictic paths while processing motion language online. Forty Chinese native speakers (NSs) and eighty English-speaking learners of L2 Chinese participated in an online judgment task. They listened to motion sentences containing deictic paths while simultaneously watching a motion display of a toward- or away-direction. Since simultaneous presentation of the sentence and the display of the same directionality require the same neural structures to process competing inputs, interference effects are expected and the reaction time to respond should take longer. Results of repeated measures ANOVA show interference effects for the NSs, but not for the L2 learners of both heritage and foreign language backgrounds, suggesting that while the NSs were sensitive to the deictic cues and automatically performed mental simulations of the deictic paths, the L2 learners' listening for imagery did not pattern with the NSs. The results added to our understanding of L2 learners' development of TFS in the new modality of listening for imagery.

Keywords: thinking for speaking; mental simulation; mental imagery; deictic paths; motion events; heritage language learners

1. Introduction

The extent to which language influences thinking has been a focus of study for decades in the fields of anthropology, linguistics and first language (L1) acquisition, psychology, and recently in second and bilingual acquisition research. Focusing on the influence of language on online thinking, Slobin's thinking-for-speaking hypothesis (TFS) [1–3] has provided new insights into this topic and attracted substantial research that explores issues pertaining to conceptual changes in language learning (e.g., [4–8]). This hypothesis suggests that speakers are habitually attuned to aspects of an event that are readily codable in the language while they are formulating speech. This TFS process varies cross-linguistically and can be observed in all forms of production and reception such as thinking for speaking/writing, listening/reading for understanding, thinking for translating, or listening for imaging [2,9]. Although Slobin's TFS framework underscores the online effects of language on thought processes, there have been very limited experimental methods employed to study the real time processing of TFS, especially for second language (L2) learners. It remains unclear whether or how well L2 learners can acquire the thought patterns of the L2 while processing the language online. This study filled the gap by adopting a simulation-based approach [10,11] to examine patterns of mental simulation produced by L1 speakers and L2 learners in real-time processing of motion language containing deictic paths.

1.1. Language and Thought in Motion Language Processing

Slobin's TFS hypothesis [1–3] claims that because each language provides a limited set of linguistic options to its speakers, instead of reflecting a direct representation of a situation, the activity of thinking takes a particular quality when one is engaging in the activity of speaking. He called this dynamic process "thinking for speaking", which involves attending to aspects of objects or events that are readily codable in the language when preparing, producing, and interpreting verbal messages. Following Talmy's [12,13] two-way motion event typology, Slobin [9,14] examined narratives produced by speakers of satellite-framed and verb-framed languages (S-languages and V-languages, respectively) and identified distinct characteristics between the two groups. Because of the large differentiated manner verb lexicons offered in S-languages, their speakers are habitually attuned to the fine-grained distinctions of manner, so as to capture features of manner. By contrast, V-languages do not encode manner in the main verbs but require their speakers to encode the path information. While V-languages also provide speakers the option of encoding manner via an adverbial or gerund outside of the main verb, this means requires more processing effort. Consequently, when describing the same motion events, speakers of S-languages encode manner information more frequently than speakers of V-languages. The same patterns were also found in mental imagery in Slobin [9]. He gave English (S-language) and Spanish (V-language) speakers the same passages to read and asked them to report mental imagery for the protagonist's movement. The passages were extracted from Spanish novels, in which manner verbs were not used but can be inferred from descriptions of the terrain and the protagonist's inner state. English speakers were given a literal translation of the Spanish texts. Interestingly, all English speakers reported mental imagery for the manner in which the protagonist moved using verbs such as *stagger*, *stumble*, or *dodge*, whereas the Spanish speakers focused on the static surroundings of the scene. Other empirical studies also confirmed that speakers of S- and V-languages attend to different components of motion events while producing or interpreting linguistic communications about motion (e.g., [6,15]).

1.2. Listening for Imagery by Native Speakers

Other methods that have been used to study processing of motion language are behavioral experiments designed to study mental imagery or simulation as well as the activation of neural patterns corresponding to perceptual or motor experiences, which are grounded in a variety of experiential domains, including cognitive, physiological, biological, and cultural (cf. [10]). Mental simulation or imagery as a basic form of cognition plays a crucial role in many thought processes such as communication, navigation, memory, and problem solving [16]. Simulation-based theories of language processing claim that understanding language involves the automatic and unconscious running of mental simulation related to the content of the utterance [10,17–23]. Studies using neuroimaging technology, such as Positron Emission Tomography (PET) and functional Magnetic Resonance Imaging (fMRI), have provided convergent evidence concerning the localization of simulation (e.g., [24–26]). In these studies, it was observed that mental simulation for understanding an utterance uses the same neural circuitry as those activated when one actually perceives or performs the action named in the utterance. That is, when a language learner runs a simulation in understanding a motion sentence such as *Throw the ball to me*, the neural motor structures responsible for the action of throwing a ball are activated as well.

Two kinds of behavioral predictions are commonly made in simulation-based approaches: compatibility effects and interference effects [10]. The distinction between these two effects lies in the timing of the presentation of stimuli. Compatibility effects are expected when the presentation of the sentence and a corresponding image/action *do not overlap* temporally, and the presentation of the sentence appears before the image or action. In such a design, the simulation that occurs while processing the preceding sentence facilitates the response in the following task if the image or action matches the sentence in terms of orientation, shape, type of action, etc. Such a facilitation effect occurs because the neural regions responsible for comprehending the content of the sentence are

activated, which primes the activation of the same neural regions for the subsequent compatible task. Zwaan et al. [23] used an experimental design to elicit compatibility effects. Participants first heard a sentence describing the motion of a ball toward or away from the listener (e.g., toward condition: *The shortstop hurled the softball at you*; away condition: *You hurled the softball at the shortstop*). They then saw pictures of objects presented in a smaller-bigger sequence (suggesting movement toward) or a bigger-smaller sequence (suggesting movement away) and were asked to decide whether the two pictures displayed the same object. A compatibility effect was observed. Participants responded more quickly to visual stimuli that matched the direction of movement described in the sentence they had heard, showing that the participants constructed a visual simulation during sentence comprehension.

The second kind of behavioral prediction is that of an interference effect. When the presentation of the sentence and the image/action *overlap* temporally, or the presentation of the sentence appears after the image or action, interference effects between the two tasks can be expected. In contrast to compatibility effects, the simultaneous presentation of the sentence and the image/action requires the same neural structures to process multiple tasks simultaneously, and thus processing is slowed for the two tasks. Kaschak et al. [11] is an example of a method that was designed to elicit interference effects. Participants were asked to decide whether the sentence they heard was meaningful or not while simultaneously watching a visual presentation of motion. The critical sentences described events involving movement in one of the four directions (up, down, toward, or away), such as *The car approached you* (toward condition) or *The car left you in the dust* (away condition), and the visual presentation also depicted one of the four directions. It was found that participants took longer to decide on the meaningfulness of the sentence when the visual stimulus and the sentence both involved motion in the same direction, because the same cognitive structure was required to process two competing tasks.

Simulations evoked by linguistic input can reflect universally-shared human experiences, as well as language-specific differences in patterns of imagery [10,27]. Such differences in the construction of mental simulations can be illustrated by the following two scenarios, involving boiling water and chewing betel nuts. Most people are familiar with the perceptual experience of seeing and hearing boiling water. When hearing the sentence *The water on the stove is boiling*, people unconsciously run a simulation depicting the heat, vapor, or bubbles rising from the bottom of the pot. However, some experiences can be specific to languages, cultures, and individuals. For instance, upon hearing the sentence *Chewing betel nuts brought him a burning sensation and made his teeth and mouth bloody red*, the mental representation evoked by this linguistic input can vary profoundly, depending on whether one has had the experience of chewing betel nuts or has seen other people do so, or whether one has other factual knowledge of betel nuts. For example, a mouth cancer expert may have a simulation with more vivid details concerning the medical effects of betel nut chewing, which would be unlikely to take place for people who lack such knowledge or relevant experience. Given that mental simulations are specific to languages and experiences, the typologically different TFS patterns observed between speakers of S-languages and V-languages are likely to be implicated in the existence of different simulation patterns between such two types of speakers.

1.3. Listening for Imagery by L2 Learners

Few attempts have been made to investigate whether and how language-driven mental imagery developed in L2 learning. Following the method used in Bergen et al. [19] with English NSs, Wheeler and Stojanovic [28] used an image-verb forced-choice matching task to test if proficient L2 English learners show the same interference effects as L1 learners. They showed first an image depicting an action and then a verb describing an action (e.g., *grab*, *kick*, or *lick*). The image and verb either used the same effector (e.g., *grab* and *push* both use hand) or different effectors (e.g., *grab* uses hand; *lick* uses mouth). The participants' task was to determine whether or not the image and the verb depicted the same action as quickly as possible. In the non-matching conditions, they found that it took the participants longer to decide when the image and verb used the same effector than when they did not.

When processing the image and verb that used the same effectors (e.g., image of *jump* followed by verb *kick*), it required co-activation of the same neural representations for processing the image and the verb, which resulted in a longer response time for two competing inputs involving the same areas of motor cortex. By contrast, when the image and verb depicted actions using different effectors, there was less similarity in the neural representations and thus made the simultaneous processing easier (cf. [29,30]). The 40 L2 participants in Wheeler and Stojanovic [28] were mostly international students studying in a public university in the United States and their mean length of English study was 14 years. Although the response times overall were longer than the NSs in Bergen et al. [19], the results suggested that proficient L2 learners can perform automatic mental simulation while comprehending motion language in a way similar to NSs.

In another study on simulation with English prepositions, Shoen [31] examined patterns of mental simulation in L1 and L2 English learners. The 36 L1 participants and 35 L2 participants first heard a sentence containing a preposition (e.g., the man went *up* the mountain) and then saw animated geometrical shapes representing the semantic content of the preposition with a moving object and a fixed landmark. The visual display either matched or mismatched (i.e., match condition: upward display vs. mismatch condition: downward display) the preposition heard in the sentence. The participants' task was to answer whether the moving object they saw was a square or a circle as fast as possible. The results suggested presence of compatibility effects for matching conditions, but was not statistically significant for the L1 group. The L2 group, by contrast, did not pattern with the L1 group and no statistically significant result was found.

While there have been a substantial number of studies on the role of mental simulation in sentence processing by L1 learners, little research has been conducted to explore L2 learners' online performance in this regard. Focusing on L1 and L2 TFS, the present study aims to explore whether or how well L2 learners of heritage and foreign language backgrounds engage in mental simulation of deictic paths while processing motion language online.

1.4. Chinese Deictic Paths

According to Talmy's motion event typology [12,13,32], the Chinese language typically uses path satellites to denote path information, which falls in the classification of a satellite-framed language. Slobin [14,33] reanalyzed such constituents as full path verbs with an equal status as the preceding manner verb, and treated both the manner verb and the path verb together as a serial verb construction. He proposed serial-verb languages such as Chinese and Thai should be classified into a third type-equipollently-framed languages. The debate over whether or not Chinese should be considered an equipollently-framed language has been a topic of heated discussion with no general consensus being reached thus far (see discussions of different stances in [32,34–39]. Nevertheless, it is generally agreed that Chinese encodes deictic paths more frequently than English. Chen [35] (p. 53) reported that 55% of the motion event descriptions found in 59 Chinese frog stories encoded deictic paths. Wu [40] examined 240 motion events that involved movement of an object produced by 40 Chinese speakers and found that 93% encoded deictic paths. Slobin [14] also commented that space deixis seems to be more closely tied to conceptions of path for Chinese speakers, as compared to the other groups of speakers.

1. a.   小   王   走   進   **來**   了。
         Xiǎo   *Wáng*   *zǒu*   *jìn*   ***lái***   *le*
         Little   Wang   walk   into   **hither**   CRS
         'Little Wang walked in [hither]'.

    b.   他   把   椅子   搬   進   辦公室   **去**   了。
         Tā   bǎ   yǐzi   bān   jìn   bàngōngshì   **qù**   le.
         he   Prep.   chair   move   into   office   **thither**   CRS
         'He moved the chair into the office [thither]'.

In addition to a path particle such as *into*, *out*, or *down* that indicates direction of a movement, Chinese speakers frequently attach a deictic hither/thither path from the perspective of the speaker: *lái* 'moving toward the speaker' or *qù* 'moving away from the speaker'. The deictic path informs the interlocutors of the relative location where one is standing and where the moving figure is heading. As shown in example (1a), the hither path *lái* denotes that the agent Little Wang is walking into a space and moving toward the speaker. Substituting the hither path *lái* with the thither path *qù* would completely change the relative location of the two, where Little Wang is walking into another space further away from the speaker. If the deictic cue is overlooked or misinterpreted, the interlocutors would not be able to quickly construct the correct relative location between oneself and the moving figure. Such use of space deixis is habitually encoded when describing movement of objects as well. For instance, the thither path *qù* in (1b) tells that the movement of the inanimate object *shū* 'book' to the office is in a direction away from the interlocutors, suggesting that they are in a location other than the office. Note that the use of deictic paths is mostly a matter of choice, but there are occasions when it becomes a matter of grammaticality. Specifically, when the sentence does not have a location or object noun (e.g., removal of 'office' in example (1b)), it becomes necessary to encode the deictic path for the utterance to be grammatical (cf. [7]).

Properly uttering deictic expressions involves several simultaneous dimensions of processing. Interlocutors need to resort to perception, proximity, and ongoing interaction in the context of utterance [41,42]. For two interlocutors to communicate effectively, they need to share not only the same grammar, but also the same pragmatically appropriate ways in orienting themselves verbally and perceptually in context. Encodings of deictic paths presents language-specific TFS patterns, in which the deictic path needs to be frequently attended to in different social situations. It has been noted that this conceptual restructuring presents considerable challenge for L2 learners whose L1 does not require such encodings of deictic relations. The deictic paths were often misused or ungrammatically omitted by English-speaking learners of L2 Chinese [7,40]. Wu [7,40] also reported that heritage language learners (HLLs) performed better than foreign language learners (FLLs) of the same proficiency level in supplying deictic paths, potentially because HLLs have more opportunities to use deictic paths in different social contexts in real-life situations. It is therefore important to examine whether there is a difference in processing deictic paths between these two types of learners.

As suggested by Slobin [1], L2 learners may require a relatively long period of time to restructure their L1 TFS in order to be able to express motion events fluently in the L2. To explore the underlying cognitive processes that may not be obvious from speech, it would be desirable to apply methods developed in simulation-based research to explore the online processing of deictic paths in an L2. If L2 learners can process Chinese deictic paths in a way similar to L1 learners, which would require them to show similar simulation patterns as the L1 learners, this would suggest that they have successfully internalized the L2 TFS patterns.

2. Research Design

The present research partially replicated the design in Kaschak et al. [11] with L1 speakers and included both L1 and L2 participants to explore the "listening for imaging" phenomenon with online processing of the hither/thither paths. Specifically, it investigated: (a) the extent to which L1 and L2 participants run mental simulations when comprehending motion sentences that contain the Chinese deictic path *lái* 'moving toward the speaker' or *qù* 'moving away from the speaker'; and (b) how such simulations might relate to participants' language backgrounds or Chinese proficiency levels. The study was approved by the Institutional Review Board at a public university in the United States. All participants provided informed consent to participate in the study prior to the data collection.

2.1. Subjects

Forty L1 and 80 L2 learners of Chinese participated in the study. The L1 participants were international students or their spouses from China or Taiwan. The L2 participants were composed of

heritage or foreign language learners enrolling in an intermediate-level Chinese language course or higher at the time when the study was conducted. To examine whether learner background plays a role in conceptual changes in processing deictic paths, the L2 participates completed a language background questionnaire to determine their status as either HLLs or FLLs. The criteria for classification of an L2 participant as an HLL were: learners (1) who identified his or her strongest language before the age of five as Mandarin Chinese or another Chinese dialect; or (2) who had one or both parents with Mandarin Chinese or another Chinese dialect as their native or dominant language, and that he or she also reported exposure to the language at home.

The L2 participants also completed a Chinese elicited imitation test (EIT) [43] to assess their L2 Chinese proficiency. The EIT was a Chinese version of the EIT developed by Ortega et al. [44], in which they tested parallel EITs in English, German, Spanish, as well as Japanese and confirmed that the EITs offered a good indication of learners' global L2 proficiency. The EIT was composed of 30 sentences ranging from 7 to 19 syllables in length and contained different vocabulary and grammatical patterns reflecting a varying degree of difficulty. The participants were required to repeat the sentence they heard as exactly as possible and their repetition was recorded and graded based on a 5-point scoring rubric for each sentence. No response or the repetition of only one word was given a score of 0, repetition of half of the sentence or less was given a score of 1, repetition that showed a change in content or form affecting the meaning was given a 2, accurate repetition of content with some changes to form was given a score of 3, and a perfect repetition was given a score of 4 (final full score = 120). A score of 50 was adopted as a cut point to divide the participants into a low or high proficiency group. Table 1 summarizes the L2 groups classified by language background (HLLs vs. FLLs) and proficiency level (Low vs. High) based on results of the EIT scores.

Table 1. L2 groups by language background and proficiency level.

Groups	N	Minimum	Maximum	Mean	SD
Low					
HLLs	14	22	47	36.21	7.04
FLLs	24	15	49	35.04	11.46
High					
HLLs	26	51	112	76.23	16.27
FLLs	16	52	115	72.00	18.40

2.2. Materials and Procedure

An online judgment task was developed to examine interference effects that arise when simultaneously processing linguistic stimuli and perceiving visual stimuli. Participants listened to Chinese motion sentences containing a deictic path *lái* or *qù* that could easily be interpreted as either toward or away from themselves (see examples (2a) and (2b)) while simultaneously watching a black-and-white motion display.

2.   a.   Toward-sentence

車子	開	過	**來**	了。
Chēzi	kāi	guò	**lái**	le.
car	drive	over	**hither**	CRS

'The car is approaching'.

       b.   Away-sentence

車子	開	過	去	了。
Chēzi	kāi	Guò	**qù**	le.
car	drive	Over	**thither**	CRS

       'The car is moving away'.

Examples (2a) and (2b) are one of the 16 pairs of critical sentences used in the experiment. In each pair, the two sentences differed only in the use of the deictic path *lái* or *qù*, which resulted in a contrast in directionality for moving either toward or away. All of the critical sentences were controlled so as to be between 6 and 9 syllables in length and were composed of vocabulary that was judged appropriate for the L2 learners' proficiency levels. A complete list of the critical toward- and away-sentences is provided in Appendix A.

The motion display (see Figure 1 for a static screenshot) was intended to convey a sense of objects moving either toward or away from the participants. The toward-display was constructed by presenting an image of a sky filled with shooting stars moving from the center to the boundaries, and the away-display was created by reversing the timeline of the presentation, which resulted in the image of stars moving from the boundaries to the center. Each motion display was presented using a resolution of 450 × 338 pixels at a rate of 15 frames per second.
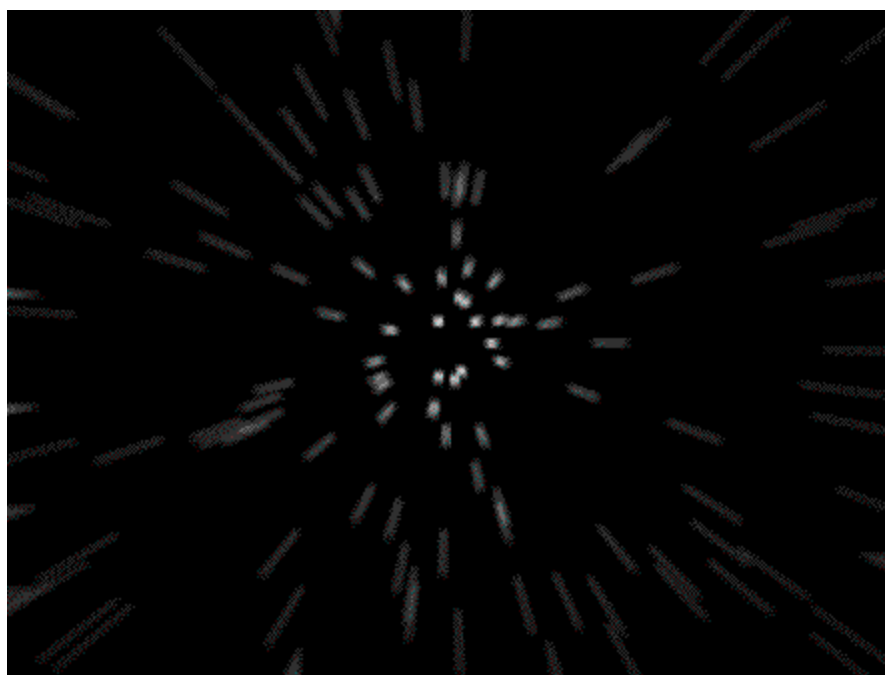

Figure 1. Visual display for toward- and away-motion.

A match condition was produced when the sentence described motion in the same direction as the motion depicted in the visual display. In contrast, a mismatch condition was generated when the sentence described motion in a direction opposite to that depicted in the display. The participants' task was to decide whether or not the sentence they heard made sense. They pressed a key labeled "Y" if the sentence made sense, or pressed a key labeled "N" if the sentence did not make sense. They were told that their responses would be timed and they should respond as quickly as possible while still maintaining accuracy.

The prediction was that participants would take longer to decide whether the sentence made sense in a match condition than in a mismatch condition. An interference effect was expected because of the simultaneous presentations of the display and the sentence of the same directionality required

the same neural structures to process the competing inputs, leading to interference between the two tasks (cf. [10,11]). For the L2 learner participants, if they were sensitive to the deictic cues and were able to simulate the toward- and away-imagery, they would show the same patterns of reaction times as the Chinese NSs in their online L2 listening-for-imagery performance.

In order to disguise the critical sentences, 24 filler sentences, including 12 sensible sentences and 12 non-sensible sentences (see examples (3a) and (3b)), were included in the experiment. The filler sentences were similar to the critical sentences in both length and level of difficulty for vocabulary.

3.　　a.　　　Sensible filler sentence

這	套	公寓	很	漂亮。
Zhè	*tào*	*gōngyù*	*hěn*	*piàoliang.*
this	CL.	apartment	very	beautiful

'This apartment is very beautiful'.

　　　b.　　　Non-sensible filler sentence

電影	去	打球	了。
Diànyǐng	*qù*	*dǎ-qiú*	*le.*
movie	go	play-ball	CRS

'The movie has left to play ball'.

Four lists were generated, with lists 1 and 2 presenting the toward-display and lists 3 and 4 presenting the away-display. The 16 pairs of critical sentences and fillers were counterbalanced across lists. Participants were randomly assigned to one of the lists.

2.3. Procedure

The experiment was conducted on a PC laptop computer using SuperLab 4.5 Stimulus Presentation Software (Cedrus Corporation: Phoenix, AZ, USA). The experiment included four blocks, with each block having 10 sentences, including 4 critical sentences (2 toward- and 2 away-sentences) and 6 fillers (3 sensible and 3 non-sensible sentences) that are presented via headphones, while simultaneously showing the toward- or away-visual display. Participants were instructed to stare at the screen. A green fixation cross appeared before each block started, and a red fixation cross appeared after the block ended. A screen asking if the participant was ready for the next block appeared between the blocks. The participant then pressed any key to continue with the next block. During each block, the 10 sentences were presented with 4-seconds interval between sentence onsets for the L1 participants and with 7-second interval for the L2 participants. The L2 participants were given more time given that L2 sentence processing can be expected to require more time than L1 processing. The length of the adjusted interval was determined through pilot tests conducted on both low- and high-proficiency L2 learners. A practice trial containing 5 sentences was given prior to the experimental blocks. The participants had to respond correctly for all five practice sentences before they were allowed to proceed to the experimental blocks. Reaction times were measured and compared between the matching and non-matching conditions and among different groups, and accuracy for each sentence was also recorded. The L1 participants spent about 5–7 min completing the experiment, and the L2 participants spent about 8–10 min.
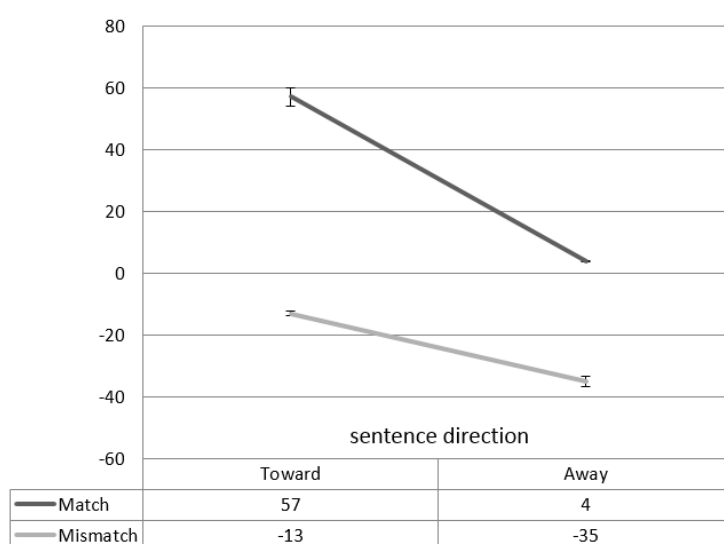
3. Results

3.1. Patterns of Simulation by L1 Participants

The reaction times were calibrated for analysis. A calibrated value of 0 ms was generated when a response occurred at the end of the sentence. A negative reaction time was generated when a response was made before the sentence ended, and a positive value when a response was made after the sentence

ended. The following procedures were adopted to eliminate outliers that could confound the results. First, participants were checked for having accuracy rate lower than 80% or having mean reaction times that were more than 2.5 standard deviations from all participants' mean reaction times. No participants were excluded due to these criteria. Next, incorrect responses were excluded. Responses that were more than 2.5 standard deviations from the mean for each participant were also removed. These two procedures led to the exclusion of less than 5% of the data. The average accuracy rate for all responses was 95.93%.

The remaining response times were submitted to a 2 (Sentence Direction: Toward vs. Away) × 2 (Compatibility: Match vs. Mismatch) repeated measures ANOVA. Figure 2 presents the mean reaction times for each condition. The effect for sentence direction was not significant, $F(1, 39) = 3.771$, $p = 0.059$, partial $\eta^2 = 0.088$, suggesting the toward- and away-sentences were generally processed at the same speed. Responses were significantly faster when the direction implied by the visual display did not match that of the sentence than when there was a match, $F(1, 39) = 6.160$, $p = 0.017$, partial $\eta^2 = 0.136$. There was no significant interaction between sentence direction and compatibility, $F(1, 39) = 0.392$, $p = 0.535$, partial $\eta^2 = 0.010$.

sentence direction		
	Toward	Away
Match	57	4
Mismatch	-13	-35

Figure 2. Mean reaction times for L1 group.

These results are in agreement with Kaschak et al. [11], supporting the predicted interference effects. When the L1 participants simultaneously processed the visual display and the motion sentence, they took longer to decide whether the sentence made sense when the visual display and the sentence depicted motion of the same direction. This suggested that language comprehension requires the same general mechanisms that are used in perception of motion. The mental simulation activated by deictic cues is thus sensitive to the directional axis of toward versus away. The NSs' automatic online construction of mental simulations for the deictic cues shows that they are tuned to the cue of deictic paths in their L1 listening-for-imagery.

3.2. Patterns of Simulation by L2 Participants

The L2 learners' responses were screened following the same outlier analysis procedures as was conducted on the L1 data. The results show that a portion of the L2 participants in each group had an accuracy rate falling below the 80% accuracy threshold. Those learners' data were thus removed, because their responses suggested that there were other confounding factors that had influenced their performance, which were beyond the focus of this study. Table 2 summarizes the results of the accuracy analysis.

Table 2. Outlier analysis: Accuracy rates for NS, HLL, and FLL participants.

Group	N	Mean Accuracy	Num. of Participants above 80%
NSs	40	96%	40
Low			
HLLs	14	58%	2
FLLs	24	61%	4
High			
HLLs	26	80%	15
FLLs	16	81%	9

As can be seen in Table 2, only two HLLs and four FLLs from the low-proficiency groups had an accuracy rate above 80%, and the mean accuracy rate for the respective group fell below the 80% threshold. Even though the learner participants were given a longer interval of 7 s to determine whether the sentence made sense, it appeared that the majority of the low-proficiency learners were unable to succeed in this task. This suggests either that the low-proficiency learners needed a longer time to process the meanings of the sentences or that the task required more advanced L2 listening skills than their current levels. On the other hand, the high-proficiency learners were more successful in this task. The mean accuracy rates for both HLLs and FLLs met the 80% threshold. However a portion of the participants were still excluded because they were below the threshold. The remaining participants from the high-proficiency groups were 15 from HLLs and nine from FLLs.

Because the numbers of the participants from the low-proficiency groups were not sufficient for an inferential analysis, the remaining participants from the two proficiency levels were then merged into just two groups, namely, HLLs and FLLs. The new HLL group comprised 17 participants and the new FLL had 13. Among these participants, no one was excluded due to having a mean reaction time that was more than 2.5 standard deviations from the group mean. The remaining responses for HLLs and FLLs were, respectively, submitted to a 2 (Sentence Direction: Toward vs. Away) × 2 (Compatibility: Match vs. Mismatch) repeated measures ANOVA to examine whether the interference effects could be found as observed in the L1 data.

The results for HLLs are presented in Figure 3. The effect for sentence direction was not significant, $F(1, 16) = 1.899$, $p = 0.187$, partial $\eta^2 = 0.106$, nor was the effect for compatibility, $F(1, 16) = 0.316$, $p = 0.582$, partial $\eta^2 = 0.019$. There was no significant interaction between sentence direction and compatibility, $F(1, 16) = 0.002$, $p = 0.969$, partial $\eta^2 = 0.000$.

	sentence direction	
	Toward	Away
Match	707	790
Mismatch	655	733
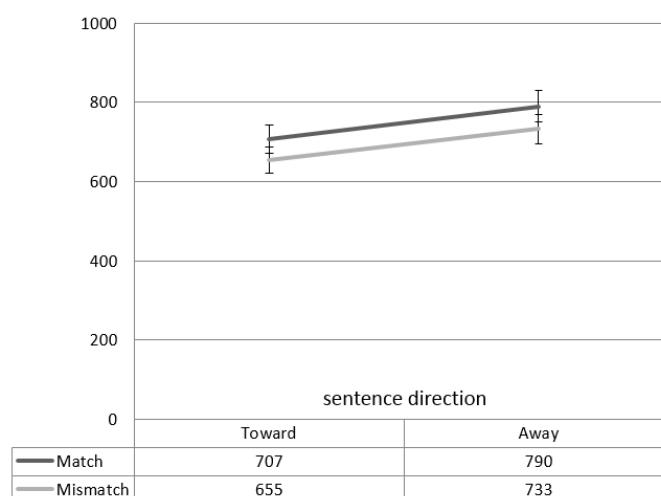
Figure 3. Mean reaction times for HLLs.

Turning to the results for FLLs, Figure 4 presents the mean reactions for each condition. The results patterned with those of HLLs. The effect for sentence direction was not significant, $F(1, 12) = 0.570$, $p = 0.465$, partial $\eta^2 = 0.045$, nor was the effect for compatibility, $F(1, 12) = 3.386$, $p = 0.091$, partial $\eta^2 = 0.220$. There was no significant interaction between sentence direction and compatibility, $F(1, 12) = 2.265$, $p = 0.158$, partial $\eta^2 = 0.159$.

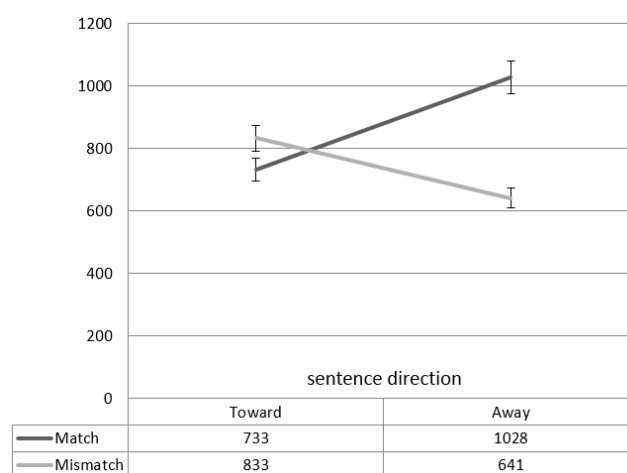| sentence direction	Toward	Away
Match	733	1028
Mismatch	833	641

Figure 4. Mean reaction times for FLLs.

4. Discussion and Conclusions

Similar to findings reported in Wheeler and Stojanovic [28] and Shoen [31], comparisons of the mean reaction times between the L1 and L2 groups showed that the L2 learners spent longer time to make a response than did the NSs. However, neither the HLLs nor the FLLs showed patterns of interference effects as observed in the L1 data, suggesting that, while these L2 learners were capable of judging whether the Chinese sentences were sensible or not, processing the deictic paths *lái* or *qù* by listening did not trigger mental imagery directed toward or away which could interfere with the simultaneous processing of the visual presentations like did the L1 group. Although the majority of the participants were from the high-proficiency groups, the L2 learners' online processes of listening-for-imagery seemed to indicate that they had not yet totally internalized the L2-specific deictic paths in their L2 TFS, even with advanced proficiency. While L2 learners were able to produce motion language following L2 TFS patterns in the mode of writing or speaking as reported in previous studies (e.g., [5,7,8]), scrutinization of the online automatic processing of the motion language in the mode of "listening for imagery" may reveal a different story. On the other hand, the results of this study were in agreement with findings reported in Yoshioka and Kellerman [45], Choi and Lantolf [46], and Stam [47], in which they also examined the co-speech gesture produced by the L2 learners and found no sign of switching to L2 TFS (cf. Gullberg [48] for discussion on gesture and bilingualism). Given that Slobin's TFS hypothesis has highlighted cross-linguistic differences in online processes for language production and comprehension, it is crucial to examine the underlying cognitive mechanisms L2 learners adopt in real time language processing. The nonverbal and online methods adopted in the present study and the aforementioned gesture studies have provided a new lens to look into the development of L2 TFS. Although English speakers naturally run mental simulations of objects moving toward them when processing motion sentences such as *The car approached you*, automatic activation of simulation in processing Chinese deictic paths *lái* 'hither' or *qù* 'thither' proved to be a domain of L2-specific TFS patterns for English-speaking learners that had not yet been salient enough to be noted when the learners were listening to the motion sentences. Echoing Wu [8,40], it was reported that L2 learners of both HLL and FLL backgrounds produced significantly less deictic paths when describing motion scenes, as compared to Chinese NSs.

Nevertheless, caution must be made when interpreting the results. The lack of statistical significance for the compatibility factor could be attributed to the smaller number of participants who contributed data to the final analysis. Moreover, it was observed that, although the L2 learners were instructed to look at the computer screen while listening to the sentences, some of them occasionally avoided looking at the screen in order to better concentrate on listening to the Chinese sentences. Hence, whether the L2 learners were capable of simulating the deictic paths online awaits further research.

The results of the study also provide some suggestions for the future design of simulation-based experiments with L2 participants. First, simulation-based methodology uses reaction times to determine whether learners are engaging in online mental simulations of the content of the sentence, but low- or intermediate-proficiency learners may not be suitable for such designs, because their limited L2 processing abilities and varied L2 listening skills are very likely to complicate the results. Additionally, a design that is meant to elicit interference effects may present more of a challenge for L2 learners than one that is meant to generate compatibility effects. This is because the processing of L2 linguistic stimuli and performance of another perceptual or motor task at the same time, as required in an interference design, consumes more attentional resources than what is required in a compatibility design, in which the two tasks do not overlap temporally. Thus, an interference-based design may require participants with more advanced L2 processing skills than that required in a compatibility design. As observed in this study, some L2 participants tended to shift their focus from one of the required tasks. It is therefore important to videotape experiment sessions to keep track of participants' actual performance.

Conflicts of Interest: The author declares no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CRS	Particle for Currently Relevant State
Disp.	Disposal Marker
Dur.	Durative Marker
Pass.	Passive Marker

Appendix A

List of Critical Sentences

1. a.

車子	很快地	衝	過	來。
Chēzi	*hěn kuài de*	*chōng*	*guò*	*lái.*
car	very quickly	rush	over	hither.

    'The car is very quickly rushing over here'.

    b.

車子	很快地	衝	過	去
Chēzi	*hěn kuài de*	*chōng*	*guò*	*qù.*
car	very quickly	rush	over	thither

    'The car is very quickly rushing away'.

2. a. 足球   飛   過   來   了。
     Zúqiú *fēi*   *guò* *lái*   *le.*
     football fly   over   hither CRS.
     'The football is flying over here'.

  b. 足球   飛   過   去   了。
     Zúqiú *fēi*   *guò* *qù*   *le.*
     football fly   over   thither CRS.
     'The football is flying away'.

3. a. 大   黃   狗   跑   過   來   了。
     Dà   *huáng* *gǒu* *pǎo* *guò* *lái*   *le.*
     big   yellow dog   run   over   hither CRS.
     'The big yellow dog is running over here'.

  b. 大   黃   狗   跑   過   去   了。
     Dà   *huáng* *gǒu* *pǎo* *guò* *qù*   *le.*
     big   yellow dog   run   over   thither CRS.
     'The big yellow dog is running away'.

4. a. 火車   開   過   來   了。
     Huchē *kāi*   *guò* *lái*   *le.*
     train   drive   over   hither CRS.
     'The train is approaching'.

  b. 火車   開   過   去   了。
     Huchē *kāi*   *guò* *qù*   *le.*
     train   drive   over   thither CRS.
     'The train is moving away'.

5. a. 老虎   衝   過   來   了。
     Lǎohǔ *chōng* *guò* *lái*   *le.*
     tiger   rush   over   hither CRS.
     'The tiger is rushing over here'.

  b. 老虎   衝   過   去   了。
     Lǎohǔ *chōng* *guò* *qù*   *le.*
     tiger   rush   over   thither CRS.
     'The tiger is rushing away'.

6. a. 籃球   被   丟   過   來   了。
     Lánqiú *bèi*   *diū*   *guò* *lái*   *le.*
     basketball Pass. throw   over   hither CRS.
     'The basketball is thrown over here'.

  b. 籃球   被   丟   過   去   了。
     Lánqiú *bèi*   *diū*   *guò* *qù*   *le.*
     basketball Pass. throw   over   hither CRS.
     'The basketball is thrown away'.

7.　a.　

他	把	球	踢	回	來	了。
Tā	*bǎ*	*qiú*	*tī*	*huí*	*lái*	*le.*
he	Disp.	ball	kick	back	hither	CRS.

'He is kicking the ball back here'.

　　b.　

他	把	球	踢	回	去	了。
Tā	*bǎ*	*qiú*	*tī*	*huí*	*qù*	*le.*
he	Disp.	ball	kick	back	thither	CRS.

'He is kicking the ball back there'.

8.　a.　

車子	開	過	來	了。
Chēzi	*kāi*	*guò*	*lái*	*le.*
car	drive	over	hither	CRS.

'The car is approaching'.

　　b.　

車子	開	過	去	了。
Chēzi	*kāi*	*guò*	*qù*	*le.*
car	drive	over	thither	CRS.

'The car is moving away'.

9.　a.　

警察	很快地	追	過	來。
Jǐngchá	*hěn kuài de*	*zhuī*	*guò*	*lái.*
police officer	very quickly	chase	over	hither

'The police officer is very quickly running over here'.

　　b.　

警察	很快地	追	過	去。
Jǐngchá	*hěn kuài de*	*zhuī*	*guò*	*qù.*
police officer	very quickly	chase	over	thither

'The police officer is very quickly running away'.

10.　a.　

球	被	打	回	來	了。
Qiú	*bèi*	*dǎ*	*huí*	*lái*	*le.*
ball	Pass.	hit	back	hither	CRS.

'The ball is hit back here'.

　　b.　

球	被	打	回	去	了。
Qiú	*bèi*	*dǎ*	*huí*	*qù*	*le.*
ball	Pass.	hit	back	thither	CRS.

'The ball is hit back there'.

11.　a.　

她	把	車	開	過	來	了。
Tā	*bǎ*	*chē*	*kāi*	*guò*	*lái*	*le.*
she	Disp.	car	drive	over	hither	CRS.

'She is driving the car over here'.

　　b.　

她	把	車	開	過	去	了。
Tā	*bǎ*	*chē*	*kāi*	*guò*	*qù*	*le.*
she	Disp.	car	drive	over	thither	CRS.

'She is driving the car away'.

12. a.
運動員	抱	著	球	跑	過	來。
Yùndòngyuán	*bào*	*zhe*	*qiú*	*pǎo*	*guò*	*lái.*
athlete	hold	dur.	ball	run	over	hither

'The athlete is holding the ball and running over here'.

b.
運動員	抱	著	球	跑	過	去。
Yùndòngyuán	*bào*	*zhe*	*qiú*	*pǎo*	*guò*	*qù.*
athlete	hold	dur.	ball	run	over	Thither

'The athlete is holding the ball and running away.'

13. a.
她	用力地	把	球	踢	過	來。
Tā	*yòng lì de*	*bǎ*	*qiú*	*tī*	*guò*	*lái.*
she	fiercely	Disp.	ball	kick	over	hither

'She is fiercely kicking the ball over here'.

b.
她	用力地	把	球	踢	過	去。
Tā	*yòng lì de*	*bǎ*	*qiú*	*tī*	*guò*	*qù.*
she	fiercely	Disp.	ball	kick	over	thither

'She is fiercely kicking the ball away'.

14. a.
他	生氣地	衝	進	來。
Tā	*shēngqì de*	*chōng*	*jìn*	*lái.*
he	angrily	rush	into	hither

'He is angrily rushing in'.

b.
他	生氣地	衝	進	去。
Tā	*shēngqì de*	*chōng*	*jìn*	*qù.*
he	angrily	rush	into	thither

'He is angrily rushing away'.

15. a.
飛機	飛	過	來	了。
Fēijī	*fēi*	*guò*	*lái*	*le.*
airplane	fly	over	hither	CRS.

'The airplane is flying over here'.

b.
飛機	飛	過	去	了。
Fēijī	*fēi*	*guò*	*qù*	*le.*
airplane	fly	over	thither	CRS.

'The airplane is flying away'.

16. a.
老鷹	快速地	飛	過	來。
Lǎoyīng	*kuài sù de*	*fēi*	*guò*	*lái.*
eagle	quickly	fly	over	hither

'The eagle is flying over here'.

b.
老鷹	快速地	飛	過	去。
Lǎoyīng	*kuài sù de*	*fēi*	*guò*	*qù.*
eagle	quickly	fly	over	thither

'The eagle is flying away'.

References

1. Slobin, D.I. From "thought and language" to "thinking for speaking". In *Rethinking Linguistic Relativity*; Gumperz, J., Levinson, S., Eds.; Cambridge University Press: Cambridge, UK, 1996; pp. 70–96.

2. Slobin, D.I. Language and thought online: Cognitive consequences of linguistic relativity. In *Language in Mind: Advances in the Study of Language and Thought*; Gentner, D., Goldin-Meadow, S., Eds.; MIT Press: Cambridge, MA, USA, 2003; pp. 157–192.

3. Slobin, D.I. Thinking for speaking. In Proceedings of the 13th Annual Meeting of the Berkeley Linguistics Society, Berkeley, CA, USA, 14–16 February 1987; Aske, J., Beery, N., Michaelis, L., Filip, H., Eds.; Berkeley Linguistics Society: Berkeley, CA, USA, 1987; pp. 435–444.

4. Brown, A. Universal Development and L1–L2 Convergence in Bilingual Construal of Manner in Speech and Gesture in Mandarin, Japanese, and English. *Mod. Lang. J.* **2015**, *99*, 66–82. [CrossRef]

5. Cadierno, T. Motion in Danish as a second language: Does the learner's L1 make a difference? In *Linguistic Relativity in Second Language Acquisition: Thinking for Speaking*; Han, Z.-H., Cadierno, T., Eds.; Multilingual Matters: Clevedon, UK, 2010; pp. 1–33.

6. Lai, V.T.; Rodriguez, G.G.; Narasimhan, B. Thinking-for-speaking in early and late bilinguals. *Biling. Lang. Cogn.* **2014**, *17*, 139–152. [CrossRef]

7. Wu, S.-L. Learning to Express Motion Events in an L2: The Case of Chinese Directional Complements. *Lang. Learn.* **2011**, *61*, 414–454. [CrossRef]

8. Wu, S.-L. Development of Thinking for Speaking: What Role Does Language Socialization Play? *Mod. Lang. J.* **2016**, *100*, 446–465. [CrossRef]

9. Slobin, D.I. Verbalized events: A dynamic approach to linguistic relativity and determinism. In *Evidence for Linguistic Relativity*; Niemeier, S., Dirven, R., Eds.; John Benjamins: Amsterdam, The Netherlands, 2000; pp. 107–138.

10. Bergen, B. Experimental methods for simulation semantics. In *Methods in Cognitive Linguistics*; Gonzalez-Marquez, M., Mittelberg, I., Coulson, S., Spivey, M.L., Eds.; John Benjamins: Amsterdam, The Netherlands, 2007; pp. 277–301.

11. Kaschak, M.P.; Madden, C.J.; Therriault, D.J.; Yaxley, R.H.; Aveyard, M.E.; Blanchard, A.A.; Zwaan, R.A. Perception of motion affects language processing. *Cognition* **2005**, *94*, B79–B89. [CrossRef] [PubMed]

12. Talmy, L. Path to Realization: A Typology of Event Conflation. In Proceedings of the 17th Annual Meeting of the Berkeley Linguistics Society, Berkeley, CA, USA, 15–18 February 1991; Sutton, L.A., Johnson, C., Shields, R., Eds.; Berkeley Linguistics Society: Berkeley, CA, USA, 1991.

13. Talmy, L. *Toward a Cognitive Semantics, Volume II: Typology and Process in Concept Structuring*; The MIT Press: Cambridge, MA, USA, 2000.

14. Slobin, D.I. The many ways to search for a frog: Linguistic typology and the expression of motion events. In *Relating Events in Narrative: Typological and Contextual Perspectives*; Strömqvist, S., Verhoeven, L., Eds.; Lawrence Erlbaum: Mahwah, NJ, USA, 2004; pp. 219–257.

15. Papafragou, A.; Hulbert, J.; Trueswell, J. Does language guide event perception? Evidence from eye movements. *Cognition* **2008**, *108*, 155–184. [CrossRef] [PubMed]

16. Kosslyn, S.M. *Ghosts in the Mind's Machine*; W.W. Norton & Company: New York, NY, USA, 1983.

17. Barsalou, L.W. Perceptual symbol systems. *Behav. Brain Sci.* **1999**, *22*, 577–609. [CrossRef] [PubMed]

18. Bergen, B.; Chang, N.; Narayan, S. Simulated Action in an Embodied Construction Grammar. In Proceedings of the 26th Annual Conference of the Cognitive Science Society, Chicago, IL, USA, 4–7 August 2005; Lawrence Erlbaum: Mahwah, NJ, USA, 2004; pp. 108–113.

19. Bergen, B.; Narayan, S.; Feldman, J. Embodied Verbal Semantics: Evidence from an Image-Verb Matching Task. In Proceedings of the 25th Annual Conference of the Cognitive Science Society, Boston, MA, USA, 31 July–3 August 2003; Lawrence Erlbaum: Mahwah, NJ, USA, 2003; pp. 139–144.

20. Stanfield, R.A.; Zwaan, R.A. The effect of implied orientation derived from verbal context on picture recognition. *Psychol. Sci.* **2001**, *121*, 153–156. [CrossRef]

21. Winawer, J.; Huk, A.; Boroditsky, L. A motion aftereffect from visual imagery of motion. *Cognition* **2010**, *114*, 276–284. [CrossRef] [PubMed]

22. Zwaan, R.A.; Stanfield, R.A.; Yaxley, R.H. Do language comprehenders routinely represent the shapes of objects? *Psychol. Sci.* **2002**, *13*, 168–171. [CrossRef] [PubMed]

23. Zwaan, R.A.; Madden, C.J.; Yaxley, R.H.; Aveyard, M. Moving words: Language comprehension produces representational motion. *Cogn. Sci.* **2004**, *28*, 611–619.

24. Shergill, S.S.; Brammer, M.; Fukuda, R.; Bullmore, E.; Amaro, E., Jr.; Murray, R.M.; McGuire, P.K. Modulation of activity in temporal cortex during generation of inner speech. *Hum. Brain Mapp.* **2002**, *16*, 219–227. [CrossRef] [PubMed]

25. Wheeler, M.E.; Petersen, S.E.; Buckner, R.L. Memory's echo: Vivid remembering reactivates sensory specific cortex. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 11125–11129. [CrossRef] [PubMed]

26. Wu, D.H.; Morganti, A.; Chatterjee, A. Neural substrates of path and manner of motion. *Neuropsychologia* **2008**, *46*, 704–713. [CrossRef] [PubMed]

27. Zwaan, R.A. Embodied cognition, perceptual symbols, and situation models. *Discourse Processes* **1999**, *28*, 81–88. [CrossRef]

28. Wheeler, K.B.; Stojanovic, D. Non-native Language Processing Engages Mental Imagery. In Proceedings of the 28th Annual Conference of the Cognitive Science Society, Vancouver, BC, Canada, 26–29 July 2006; Lawrence Erlbaum: Mahwah, NJ, USA, 2006.

29. Narayan, S.; Benjamin, B.; Weinberg, Z. Embodied Verbal Semantics: Evidence from a lexical matching task. In Proceedings of the 30th Annual Meeting of the Berkeley Linguistics Society, Berkeley, CA, USA, 13–16 February 2004.

30. Pulvermüller, F.; Härle, M.; Hummel, F. Walking or talking? Behavioral and neuropsychological correlates of action verb processing. *Brain Lang.* **2001**, *78*, 143–168. [CrossRef] [PubMed]

31. Shoen, B. Understanding English Prepositions: Exploring Mental Simulation in Learners of English as a Second Language. Master's Thesis, University of Hawai'i at Mānoa, Honolulu, HI, USA, Unpublished Work, 2006.

32. Talmy, L. Main Verb Properties. *Int. J. Cogn. Linguist.* **2012**, *3*, 1–23.

33. Slobin, D.I. What makes manner of motion salient? Explorations in linguistic typology, discourse, and cognition. In *Space in Languages: Linguistic Systems and Cognitive Categories*; Hickmann, M., Robert, S., Eds.; John Benjamins: Amsterdam, The Netherlands, 2006; pp. 59–81.

34. Beavers, J.; Levin, B.; Tham, S.W. The typology of motion expressions revisited. *J. Linguist.* **2010**, *46*, 331–377. [CrossRef]

35. Chen, L. *The Acquisition and Use of Motion Event Expressions in Chinese*; Lincom Europa: Munich, Germany, 2007.

36. Chen, L.; Guo, J. Motion events in Chinese novels: Evidence for an equipollently-framed language. *J. Pragmat.* **2009**, *41*, 1749–1766. [CrossRef]

37. Li, T.F. An Empirical Study on Stationary Motion Events in Mandarin Chinese. *Int. J. Cogn. Linguist.* **2013**, *4*, 133–156.

38. Peyraube, A. Motion events in Chinese: Diachronic study of directional complements. In *Space in Languages: Linguistic Systems and Cognitive Categories*; Hickmann, M., Robert, S., Eds.; John Benjamins: Amsterdam, The Netherlands, 2006; pp. 121–135.

39. Yi, J.; Hendriks, H.; Hickmann, M. The expression of caused motion events in Chinese and in English: Some typological issues. *Linguistics* **2011**, *49*, 1041–1076.

40. Wu, S.-L. Influence of L1 thinking for speaking on Use of an L2: The Case of Path Expressions by English-Speaking Learners of Chinese. In *Studies in Second Language Acquisition of Chinese*; Han, Z.-H., Ed.; Multilingual Matters: Bristol, UK, 2014; pp. 1–29.

41. Hanks, W.F. Language form and communicative practices. In *Rethinking Linguistic Relativity*; Gumperz, J.J., Levinson, S.C., Eds.; Cambridge University Press: Cambridge, UK, 1996; pp. 232–270.

42. Hanks, W.F. Fieldwork on deixis. *J. Pragmat.* **2009**, *41*, 10–24. [CrossRef]

43. Wu, S.-L.; Ortega, L. Measuring global oral proficiency in SLA research: A new elicited imitation test of L2 Chinese. *Foreign Lang. Ann.* **2013**, *46*, 680–704. [CrossRef]

44. Ortega, L.; Iwashita, N.; Norris, J.M.; Rabie, S. *An Investigation of Elicited Imitation Tasks in Crosslinguistic SLA Research*; Second Language Research Forum: Toronto, ON, Canada, 2002.

45. Yoshioka, K.; Kellerman, E. Gesture introduction of Ground reference in L2 narrative discourse. *Int. Rev. Appl. Linguist.* **2006**, *44*, 173–195. [CrossRef]

46. Choi, S.; Lantolf, J.P. REPRESENTATION AND EMBODIMENT OF MEANING IN L2 COMMUNCATION: Motion Events in the Speech and Gesture of Advanced L2 Korean and L2 English Speakers. *Stud. Second Lang. Acquis.* **2008**, *30*, 191–224. [CrossRef]

47. Stam, G. Thinking for speaking about motion: L1 and L2 speech and gesture. *Int. Rev. Appl. Linguist.* **2006**, *44*, 143–169. [CrossRef]

48. Gullberg, M. Bilingualism and gesture. In *The Handbook of Bilingualism and Multilingualism*, 2nd ed.; Bhatia, T.K., Ritchie, W.C., Eds.; Wiley-Blackwell: Hoboken, NJ, USA, 2012; pp. 417–437.