Abstract
In the current study, hand motions captured by a mouse-tracking system were used to index listener’s cognitive processes while making commitments to different choice alternatives during the processing of English passive and active structures. Fifty-seven second language (L2) speakers and 19 first language (L1) speakers of English carried out an aural forced-choice picture identification task comprised of 75 items. The findings indicate that although L1 participants have quicker response times for both active and passive structures than L2 participants, both L1 and L2 participants demonstrate similar difficulties in processing passive constructions.

Keywords: Syntactic processing; second language acquisition, embodied cognition.

Introduction
Assessing syntactic processing in first language (L1) and second language (L2) speakers is an important approach to understanding how speakers cognitively structure language. The majority of previous research on language processing and cognition has focused on off-line or indirect assessments of language processing that primarily investigate language processing as a lateral mechanism that involves a discrete-symbolic architecture. In contrast to this approach, we use an embodied cognition experimental design to compares how L1 and L2 speakers process syntactic structures. Such an approach allows us to examine how cognitive structures in L1 and L2 speakers act in parallel and evolve over short time spans.

Traditional theories of cognition have viewed motor processing as the end-result of cognitive processing. However, recent research demonstrates that action dynamics are not the aftermath of cognition but rather a requisite of cognition (Anderson, 2003). Thus, tracking the dynamics of body movement can provide evidence for cognitive processing. In the current study, we used hand motions captured by a mouse-tracking system to index listener’s cognitive processes while making commitments to different choice alternatives during the processing of English passive and active structures. The main purpose of the study was to compare processing of passive and active constructions between L1 and L2 speakers of English by using traditional on-line mechanisms such as response time in conjunction with newer on-line mechanisms that measure motor movement. We presume that passives will take longer to process and demonstrate differences in motor responses for both L1 and L2 speakers and that this effect will be greater for L2 speakers.

Language and Cognition
In traditional research, language has often been viewed as a discrete-symbolic architecture that contains elements of representation, formalism, and rule-based transformations. (i.e., the manipulation of symbols following explicit rules; Chomsky,1965). Other language learning theories such as usage-based approaches hold that linguistic patterns in the input (i.e., form-meaning mappings) afford acquisition (Ellis, 2012). One problem with such approaches is that they lack a device to ground language representations in the physical world.

The need to physically ground representations implies that there is more to cognition than only mental and linguistic representations of knowledge. Because cognition involves repeated interplay with the environment, cognition itself is theorized to be grounded (Anderson, 2003) and to integrate a variety of information sources in parallel (Freeman, Dale, & Farmer, 2011). Such theories fall under the umbrella term of embodied cognition. Embodied cognition is based on the notion that cognition exploits environmental interactions to simplify and advance cognitive tasks (Anderson, 2003; Glenberg & Robertson, 1997; Lakoff, 1987). One approach to investigate language processing is through an embodied cognition approach, which examines motor responses such as hand-motions as a reaction to language stimuli. Such responses have been shown to be continuously and temporally updated by perceptual and cognitive processing such as that found in language processing (Freeman & Ambady, 2012; Tipper, Tiffaney, 2013; Gendron, 2015; Bub, 2013).

**Language Processing and Embodied Cognition**

Automatic language processing can be elicited through various experimental methods. Generally, these methods prompt participants to respond to language stimuli as quickly as possible using response time (RT) measures (see Segalowitz and Trofimovich, 2012 for review). Motor responses, especially those found in hand motions, can also be an important component of measuring cognitive processing, especially language processing. These responses not only provide an accessible method for investigating cognitive processing, but they are unlike the majority of traditional language research methods found in L1 and L2 research which focus on off-line or indirect observations (Marinis, 2003). Motor responses also differ from traditional and more recent on-line methods such as measuring RTs or collecting eye-tracking data because they can provide strong evidence that language processing is continuous and dynamic as well as evidence that it occurs at multiple levels in parallel (Freeman et al., 2011).

Recent studies have successfully used mouse-tracking experiments to examine participants’ processing of linguistic input. For instance, Spivey, Grosjean, and Knoblich (2005) used mouse-tracking software to assess phonological awareness, finding that spoken words activate multiple lexical terms while, concurrently, the language processing mechanism continually updated the phonetic representation of the word. Recent studies using mouse-tracking technology to examine syntactic processing have been undertaken by Farmer, Anderson, and Spivy (2007) and Dale and Duran (2011). These studies support the notion that partially active syntactic constructions compete with each other over time, that constructions are influenced by visual, contextual, and linguistic factors (Farmer et al., 2007) and that constructions can involve rapid shifts in cognitive dynamics (Dale & Duran, 2011).

**Syntactic Processing: English Passives**

Our interest in this study is to assess the potential for motor responses to provide insight into the processing of passive structures by L1 and L2 speakers of English. The target structures of our study are English passive and active constructions. The development of passives in L1 and L2 speakers has been explored in previous studies from a variety of perspectives because the form and meaning mapping in passive constructions is a complex phenomenon. In passive constructions, the patient serves the grammatical subject followed by auxiliary be, a lexical verb in the past participle form, and optional by-phrase with the agent. In a passive construction, the patient role is mapped to sentence subject. On the other hand, in active constructions, subject (agent) + verb constructions follow the regular syntactic order of English. In this case, the agent role is mapped to the sentence subject. For instance, a picture of a cat scratching a chair can be described using an active sentence (e.g., A cat scratches a chair.), whereas the same message can be delivered using a passive sentence (e.g., A chair is scratched by a cat.). Although they intend to express the same meaning, the two constructions involve different mappings of thematic roles to grammatical functions and different constituent structures. Two possible reasons why speakers face challenges with processing passives compared to actives are because passives involve more complex constituent structure (i.e., the inclusion of an additional auxiliary verb and a by-phrase) and they require non-canonical mapping of thematic roles such as agents (Messinger, Branigan, & McLean, 2012).

Previous researchers in the domain of L1 and L2 acquisition research have used structural priming approaches to examine the processing and production of passive structures. Bencini and Valian (2008), for example, examined the effect of structural priming on children’s comprehension of passives. They found that while passive priming led to greater production of passive constructions than exposure to active constructions, it did not facilitate greater comprehension of passive constructions. Other studies have also examined structural priming in children and found that by age nine, children have mastered both the syntactic and thematic dimensions of passives (Marchman, Bates, Burkardt & Good, 1991; Messener, Branigan, and McLean, 2012). For L2 learners, research shows that syntactic priming facilitates the production of passives (Kim & McDonough, 2008). In sum, although research has examined the processing and development of passives, few researchers have investigated how adult L1 and L2 speakers of English process passive structures compared to active structures, especially from an embodied cognition perspective.

**Method**

The purpose of the study is to examine how tracking the dynamics of body movement can provide evidence for cognitive processing in L1 and L2 speakers of English. We used hand motions captured by a mouse-tracking system to continuously index listener’s commitments to different choice alternatives during the processing of English passive and active structures. Such an approach allows us to use online data collection to investigate language processing using an action-dynamics approach (tracking the dynamics of body movement). Furthermore, we examined whether L2 speakers’ English proficiency impacts their processing of active and passive constructions. The current study was guided by the following research question: Are there any differences in the processing of active and passive constructions between native and non-native speakers of English with two proficiency levels?

**Participants**

A total of 57 non-native speakers (NNS) of English (24 females and 33 males) and 19 native speakers (NS) of English participated in the study. Of those NS participants that completed the post-experiment survey, 10 were female and 4 were male. All NS participants were undergraduate
students who were enrolled at a major southeastern university in the US and received class credit in a freshman Psychology course for participating in the experiment. All participants had normal or corrected to normal vision. NS participants ranged in age from 19 to 33 and had an average grade point average of 3.37 (for those that completed the post-experiment survey). All NNS participants were native speakers of Spanish and were enrolled at the Instituto Tecnológico y de Estudios Superiores de Monterrey (ITESM) campus in San Luis Potosi, Mexico studying at either the high school or college level. While Spanish does have a passive structure that corresponds to English, passives are less common in Spanish than in English because Spanish has a range of structures available to mystify the agent (Blanco-Gomez, 2002). All NNS participants had normal or corrected to normal vision. The NNS participants ranged in age from 15 to 24. All NNS participants had taken a paper-based institutional TOEFL one month before the data collection. The average paper-based TOEFL score for the participants was 519. Based on the TOEFL scores, NNS participants were further divided into two proficiency groups: low and high (see the results section).

**Materials and Study Design**

For this experiment, we used a within subjects comparison design with an aural forced-choice picture identification task. The task contained 75 items in the form of complete sentences: 30 target items (15 passive constructions and 15 active constructions) and 45 distractors (15 relative clause constructions, 15 dative constructions, and 15 prepositional phrases of location constructions) (see Appendix A for the stimuli list). Since our main interest is comparing passive to active constructions, the relative, dative, and prepositional constructions were treated as filler constructions. The verbs and nouns used in the passive and active items were checked for occurrence on the General Service List (West, 1953; adapted by Baumann & Culligan, 1995) to increase the likelihood that the learners would be familiar with their meaning and use (http://www.newgeneralservicelist.org/). In order to ensure participants’ familiarity of words, we gave vocabulary tests using the list of nouns and verbs in the language stimuli to a similar group of NNS participants. Any words that were found unfamiliar were eliminated.

For each item, participants listened to a sentence (e.g., *The boy is pushed by his sister.*) and had to select the picture that corresponded to its correct meaning from two pictures. Participants could begin moving the mouse at the onset of the sentence reading. The passive and active targets involved pictures of reversible events. For example, the passive construction *the boy is pushed by his sister* was paired with pictures of a girl pushing a boy and a boy pushing a girl. Similarly, the active construction *the bus hits the motorcycle* was paired with pictures of a bus hitting a motorcycle and a motorcycle hitting a bus. All pictures were piloted with similar groups of NNS participants, and the stimuli which caused any confusion were modified.

**Apparatus and Procedure**

We used MouseTracker software (Freeman & Ambady, 2010) to collect hand motion data. MouseTracker continuously catalogs participants’ commitments to two choice alternatives during a behavioral response to language stimuli. The hand motion data can provide real-time traces of the mind’s cognitive processes including language processing (Freeman, Dale, & Farmer, 2011). From MouseTracker, we specifically collected response time (RT) data, accuracy data, and motor response data related to maximum deviation (MD), and area under the curve (AUC) for correct decisions. RTs in milliseconds were collected to assess the time it took for participants to make decisions. Accuracy data recorded whether or not the participants’ decisions were correct (a dichotomous yes/no decision). MD and AUC data examine whether mouse trajectories for one condition travel closer to an unselected alternative relative to another condition (i.e., the spatial attraction of the alternative choice). For both MD and AUC measures, MouseTracker first computes an idealized response trajectory (a straight line from the start to the endpoint for the correct selection). MD is calculated as the largest perpendicular deviation between the idealized trajectory and the actual trajectory while the AUC is the geometric area between the actual trajectory and the idealized trajectory (Freeman & Ambady, 2010).

Participants were first given instructions on how to interact with the software and told they were free to use either their left or their right-hand. They were then given eight practice trials to familiarize themselves with the task on a computer. Each trial contained a start button at the bottom center of the screen and a picture in the upper left and upper right of the screen. When the participants clicked the start button, a sentence that matched one of the pictures was presented aurally. Participants then moved the mouse to choose the picture that they thought best represented the sentence they heard. Once the mouse reached the picture, the trial stopped. Participants were asked to begin mouse movements early and were warned if their mouse movements started 1000 ms after onset of the stimuli. If a response was not made within 2000 ms, the trial was discarded. Following training, the participants were tested on the 75 stimuli in the stimuli list (the 15 passive, active, relative clause, dative, and prepositional constructions). The presentation of the stimuli and the presentation of the pictures were randomized and counterbalanced across participants. Display resolution was set to 1280 x 800.the last two decades, the scope of L2 processing research

**Data Analysis**

The data was first checked to assess participant response accuracy in the experiment to assess randomness. We set a cut off for inclusion into the analysis of 75% accuracy in selections across the passive and active structures (based on a Chi-square). This led to the removal of 11 NNS participants and three NS participants. We also assessed item difficulty for the passive and active structures for NNS participants and three NS participants. We also assessed item difficulty for the passive and active structures for NNS participants and three NS participants.
and NS participants separately. We set accuracy for individual items at 75% or greater for inclusion. For the NNS participants, this led to the removal of one item from the passive stimuli. For the NS participants, this led to the removal of three items from the passive stimuli. The remaining data were then transformed using logarithmic formulas (logarithms to the base of 10; Field, 2005; Larson-Hall & Herrington, 2010) to correct for outliers. Those outliers that remained after transformation (as identified in box plots) were removed.

For the final data, the NNS participants were split into high (n = 25) and low (n= 21) proficiency categories based on the mean TOEFL score for the group (M = 519). Within-subjects analyses of variance (ANOVAs) were used to investigate main effects between syntactic structure and participants’ response times, MD, and AUC and interactions based on language proficiency (low level NNS, high level NNS, and NS) for correct answers only. Multivariate analyses of variance (MANOVAs) were used to assess differences in active and passive structures among language proficiencies for response times, MD, and AUC for correct answers only.

### Results

**Response Time (RT)**

In order to assess speed of processing active and passive constructions, we examined the participants’ RTs. First, a within-subject ANOVA was conducted to determine the effect of syntactic structure on RT. The results showed that there was a significant main effect of syntactic structure on participant response time, $F(1, 59) = 166.630$, $p < .001$, $\eta_p^2 = .688$, indicating that active structures had a faster RT than passive structures. For RTs, there was a significant interaction between syntactic structure and language proficiency, $F(1, 59) = 5.059$, $p < .010$, $\eta_p^2 = .146$, indicating differences in RTs based on language proficiency. Pairwise comparisons revealed that low proficiency and high proficiency NNS participants’ overall response times were significantly slower than NS participants’ response times ($p < .050$).

In order to examine whether participants’ proficiency level played a role in their processing of two different structures, a between-subjects MANOVA was conducted. The results reported a significant differences between participant language proficiency levels for active structures, $F(2, 59) = 7.306$, $p < .010$, $\eta_p^2 = .199$, and for passive structures, $F(2, 59) = 3.433$, $p < .050$, $\eta_p^2 = .097$. Pairwise comparisons demonstrated that both low and high proficiency NNS participants had significantly slower RTs than NS participants for active structures ($p < .001$), but were not significantly different than one another. Pairwise comparisons also demonstrated that low proficiency NNS participants had slower RTs than high proficiency NNS participants and NS participants ($p < .050$) for passive structures, but that high proficiency NNS participants and NS participants did not statistically differ in the RT times for passive structures (see Table 1 for descriptive statistics for these results).

**Maximum Deviation (MD)**

In order to measure the maximum attraction toward the unselected alternative (i.e., the attraction to the active interpretation of a passive sentence), we examined the maximum deviation of the curve. There was a significant main effect of syntactic structure on participants’ MD, $F(1, 59) = 22.078$, $p < .001$, $\eta_p^2 = .272$, indicating that mouse trajectories deviated toward the unselected alternative to a greater degree in the passive structures as compared to the active structures. For MD, there was not a significant interaction between syntactic structure and language proficiency, $F(1, 59) = 0.257$, $p > .050$. Pairwise comparisons revealed no differences between low proficiency NNS participants, high proficiency NNS participants, and NS participants overall MDs.

A between-subjects MANOVA reported no significant differences between participant language proficiency levels for active structures, $F(2, 59) = 0.188$, $p > .050$ and for passive structures, $F(2, 59) = 0.042$, $p > .050$. Pairwise comparisons revealed no differences between low proficiency NNS participants, high proficiency NNS participants, and NS participants for either active or passive MDs (see Table 1 for descriptive statistics for these results).

**Area Under the Curve (AUC)**

In order to measure the overall attraction toward the unselected alternative, we examined the AUC. There was a significant main effect of syntactic structure on participants’ AUC, $F(1, 59) = 17.019$, $p < .001$, $\eta_p^2 = .224$, indicating that all participants had a greater overall attraction to the unselected alternative in the passive structures as compared to the active structures. For AUC, there was not a significant interaction between syntactic structure and language proficiency, $F(1, 59) = 0.955$, $p > .050$. Pairwise comparisons revealed no differences between low

### Table 1

**Descriptive statistics for active and passive responses: Mean (standard deviation)**

<table>
<thead>
<tr>
<th>Index</th>
<th>Structure</th>
<th>Low proficiency NNS (n = 21)</th>
<th>High proficiency NNS (n = 25)</th>
<th>NS (n = 16)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response time logarithmic</td>
<td>Active</td>
<td>3.412 (0.087)</td>
<td>3.373 (0.097)</td>
<td>3.298 (0.085)</td>
</tr>
<tr>
<td></td>
<td>Passive</td>
<td>3.461 (0.089)</td>
<td>3.430 (0.090)</td>
<td>3.386 (0.076)</td>
</tr>
<tr>
<td>Maximum Deviation</td>
<td>Active</td>
<td>0.417 (0.020)</td>
<td>0.430 (0.218)</td>
<td>0.392 (0.145)</td>
</tr>
<tr>
<td></td>
<td>Passive</td>
<td>0.474 (0.207)</td>
<td>0.487 (0.211)</td>
<td>0.471 (0.162)</td>
</tr>
<tr>
<td>Area under the Curve</td>
<td>Active</td>
<td>0.647 (0.515)</td>
<td>0.969 (1.039)</td>
<td>0.558 (0.339)</td>
</tr>
<tr>
<td></td>
<td>Passive</td>
<td>0.741 (0.552)</td>
<td>1.143 (1.160)</td>
<td>0.795 (0.579)</td>
</tr>
</tbody>
</table>
proficiency NNS participants’, high proficiency NNS participants’, and NS participants’ overall AUC.

A between subjects MANOVA reported no significant differences in the AUC between participant language proficiency levels for active structures, $F(2, 59) = 1.802, p > .050$ and for passive structures, $F(2, 59) = 1.479, p > .050$. Pairwise comparisons revealed no differences between low proficiency NNS participants, high proficiency NNS participants, and NS participants for either active or passive AUC (see Table 1 for descriptive statistics for these results).

**Discussion**

In the current study, we demonstrated how passive constructions in English are processed differently than active structures among NS and NNS participants using RTs and hand motions captured by a mouse-tracking system. The system continuously indexed listener’s commitments to different choice alternatives during the processing of English passive and active structures. This is a novel approach that affords the collection of information that cannot be obtained using traditional behavioral and assessment measures such as RTs and accuracy on comprehension tests, which have been predominantly used in psycholinguistic oriented second language acquisition (SLA) research.

To summarize the findings, as reported in our within-subjects ANOVAs, our participants processed active constructions faster than passive constructions based on their RTs, which supports previous research that has demonstrated the challenges of mapping thematic roles in passive constructions (Messenger, Granigan, & McLean, 2012). Additionally, NNS RTs were slower than NS RTs for both constructions regardless of NNS proficiency level. Results also showed that there were main effects for construction types on participants’ MDs and AUCs, demonstrating that mouse trajectories for both L1 and L2 participants deviated toward the unselected alternative to a greater degree in the passive as compared to the active structures. This suggests that both NS and NNS participants performed similarly.

Our MANOVA analyses looked at between-subjects differences for language proficiency levels for both active and passive constructions. Active construction RTs were slower for both low and high proficiency NNS as compared to NS, but no difference was found between the two NNS groups. For passives, low proficiency NNS were slower than high proficiency NNS and NS, but high proficiency NNS and NS were not different. There were no significant differences in either MD or AUC based on condition between low proficiency NNS, high proficiency NNS, and NS.

**Response Time Results**

The RT results suggest that passive constructions take longer to process than active constructions. Such a result likely indicates that participants had a more difficult time defining the agent and patient in the sentence because of the passive transformation that places the object noun phrase in the subject position. This finding is in line with previous L1 and L2 research on passive constructions (e.g., Bencini & Valian, 2008). The RT results also indicate that NNS at both the high and low proficiency levels were slower to respond to active constructions when compared to NS. This finding likely demonstrates that NNS have yet to reach the processing fluency of NS. However, in terms of passive constructions, low proficiency NNS demonstrated slower RTs than high proficiency NNS and NS. Tangential support for this finding can be found in Kim and McDonough (2008). Their research indicated that high-level L2 learners produced more passives during structural priming activities than low-level L2 learners. Thus, it appears that there may be differences between high and low L2 learners in both passive priming and processing. High proficiency NNS and NS did not exhibit differences in RTs for the passive constructions. This may indicate that passive constructions impede automatic processing by NS, slowing down processing to such a degree that NS RTs are similar to that of high proficiency NNS (i.e., passives slow down processing in both groups to an equal degree).

**Maximum Deviation and Area Under the Curve Results**

MD and AUC measures, which examine the spatial attraction of the alternative choice (in this case, initially processing a passive sentence as an active sentence), indicated that passive constructions interrupt cognitive activity in language processing. The results show that mouse trajectories for passive constructions deviated toward the unselected alternative (i.e., the active structure) to a greater degree than the unselected alternative in active structures (i.e., the passive structure). This finding indicates that participants initially process the object noun phrase in the passive construction as the subject noun phrase. For instance, in the sentence, *the boy is pushed by his sister*, participants initially move the mouse toward the incorrect picture that shows the boy pushing his sister and not the correct picture that shows the sister pushing the boy. Such findings provide evidence for the inherent difficulty in processing passive constructions as compared to active constructions. Intriguingly, we see no differences in either the MD or AUC results as a function of language proficiency. Thus, regardless of proficiency level with English, passive constructions cause cognitive processing difficulty in terms of comprehension. It may be the case that similar mappings between English and Spanish for the passive construction allow Spanish speakers to process passives in a manner similar to that of native speakers (in terms of motor responses). However, since passives are less frequent in Spanish than in English, recognizing and responding to passive constructions may take longer for Spanish participants (as evidenced by the differences in RTs between NS and NNS participants) because the they have not adapted to the statistical regularities found in English input (Fine, Jaeger, Farmer, & Qian, 2013).
Conclusion
The findings of the current study provide important methodological and theoretical implications for language processing and acquisition by examining the processing of English passives by L2 low and high-level learners of English and L1 speakers of English. Taking a novel approach to explore language processing (i.e., a dynamic, embodied cognition approach), the current study provided additional insights into the processing of passive and active constructions and how hand motions can be used to index listeners’ cognitive processes. Specifically, the findings demonstrate differences and similarities between L1 and L2 participants such that L1 participants are faster at responding to passive and active stimuli when compared to L1 participants. However, both groups of participants show similar difficulties in processing passive constructions as compared to active structures.

The limitations of the current study include the sample size, the focus on a single pool of L2 learners, and the lack of concurrent validity to support the findings and their interpretations. Thus, future research is warranted to investigate language processing with a larger population of L2 learners from a variety of L1 backgrounds. Future research using other online processing methods should also be conducted to provide concurrent validity for this study. Such studies could use event-related brain potential measures to provide extra information about real time processing of passive and active constructions or use eye-tracking data to investigate syntactic processing. In addition, future research using hand tracking methods to focus on a variety of other syntactic structures in various target languages that address learner variables such as language aptitude, language analytic skills, and working memory would help provide support for theories of syntactic processing and acquisition in L1 and L2 populations.

References