Embodied Language Learning in Virtual Reality

by

Christian David Vázquez Machado

Submitted to the program in Media Arts and Sciences
School of Architecture and Planning, in partial fulfillment of the
requirements for the degree of Master of Science in Media Arts and Sciences
at the Massachusetts Institute of Technology.

June 2018
© 2018 Massachusetts Institute of Technology. All rights reserved.
Embodied Language Learning in Virtual Reality

by
Christian David Vázquez Machado

Submitted to the program in Media Arts and Sciences
School of Architecture and Planning, in partial fulfillment of the
requirements for the degree of Master of Science in Media Arts and Sciences
at the Massachusetts Institute of Technology.

Abstract

Embodied theories of language propose that the way we communicate verbally is grounded in our body. Nevertheless, the way a second language is conventionally taught does not capitalize on embodied modalities. The tracking and immersive capabilities of virtual reality systems can enable a change in the way students learn language by engaging them in kinesthetic activities that explicitly use body movement to encode knowledge. The body can also be used implicitly to alter a student's perception of themselves in order to enhance the way they approach learning in immersive environments. In this work, we seek to explore the potential of both explicit and implicit embodied language learning using virtual reality as a platform. For the purpose of this thesis we focus on vocabulary acquisition to assess the potential impact these methodologies can have on language education. Two systems were developed that afford explicit (Words in Motion) and implicit (Inner Child) embodied learning. Both systems were evaluated separately during controlled experiments with 60 participants each. Explicit embodied learners displayed enhanced retention positively correlated with performing actions in the Words in Motion platform. Our findings from the implicit embodied study highlight the importance of having a body in virtual reality. Inner Child successfully increased word retention when inducing a subjective age reduction that correlated with the feeling of ownership of a virtual child avatar. These results support the hypothesis that virtual reality can deeply impact language learning by leveraging the body explicitly and implicitly.

Thesis Supervisor: Pattie Maes
Title: Professor in Media Arts and Sciences
Embodied Language Learning in Virtual Reality

by
Christian David Vázquez Machado

The following people kindly served as readers for my thesis:

Signature redacted

Thesis Reader
Mitchel Resnick
Professor of Media Arts and Sciences
MIT Media Lab

Signature redacted

Thesis Reader
Mel Slater
Distinguished Investigator
University of Barcelona
Table of Contents

Chapter 1. Introduction
   1.1 Initial Remarks 10
   1.2 Motivation 11
   1.3 Contribution 13
   1.4 Thesis Overview 14

Chapter 2. To the Mind, Through the Body
   2.1 Embodied Cognition and Language 17
   2.2 Embodied Learning 19
      2.2.1 Explicit Embodied Learning 19
      2.2.2 Implicit Embodied Learning 23
   2.3 Virtual Reality Affordances for Embodied Learning 25
      2.3.1 Fine-Grained Body Tracking 25
      2.3.2 Contextual Affinity 26
      2.3.3 Enhanced Action/Representation Space 26
      2.3.4 Becoming Others 27

Chapter 3. Words in Motion
   3.1 Explicit Embodied Language Learning 28
      3.1.1 No Gestural Congruency, No Veridical Convergence 29
      3.1.2 Gestural Congruency, No Veridical Convergence 29
      3.1.3 Gestural Congruency and Veridical Convergence 30
   3.2 System Implementation 31
      3.2.1 Recognizing Actions 31
      3.2.2 Training New Actions 32
      3.2.3 Challenges 32
   3.3 Learning Activity Design 34
   3.4 Activity Trials 35

Chapter 4. Explicit Embodied Learning Study
   4.1 Vocabulary Acquisition Experiment Design 37
      4.1.1 Word Selection 37
      4.1.2 Evaluation Method 37
Chapter 1
INTRODUCTION

1.1 Initial Remarks

Learning a second language is a journey that typically begins in the classroom. However, it is not until we get out into the world that we validate our success (or failure) at learning by employing language to accomplish our goals. This transition, or rather transfer of knowledge, from a realm of conventional (curriculum-ready) interactions to an unpredictable reality comes with its own challenges. The language-learning classroom is an environment where we can feel safe to make mistakes when we communicate with other learners. The real world, on the other hand, is not nearly as forgiving. Expressing oneself becomes a challenge of finding the right words to transmit the right set of ideas in ever-changing situations. In a way, there is a disconnect between the context in which people learn, and the way that knowledge is catalyzed in real-life scenarios. (In fact, this is not an exclusive dilemma to language learners, but to learners in general.)

Second language education has often made use of technology to enhance the way we learn, finding new ways to bridge this transfer gap and change the way language is taught [1, 2]. Due to the ubiquity of cheap and capable mobile devices, schools have shown an increase in tablet and smartphone use in the classroom to complement or augment traditional curriculums [3]. These devices are often incorporated in interactive activities that present material to the student and engages them to tackle different aspects of language learning individually or collaboratively. However, like most traditional language curriculums, the use of technology in the classroom is primarily audio-visual in nature [4, 5]. Many studies have shown that there is a direct relationship between our body and the way we encode information in our brains, making the use of the body as a tool for learning a promising approach [6]. This, in itself, is not an argument against audio-visual based learning, but it highlights a missed opportunity on behalf of technology enhanced language learning, which is not leveraging the potential of the body to its fullest.
Mixed reality (MR), a spectrum of technologies that includes augmented reality on one end and virtual reality on the other, offers a unique opportunity for language learning that breaks this paradigm. One of its greatest advantages as a tool for language education is its ability to offer rich context to learners. Be it their real environment in augmented reality or a completely immersive scenario in virtual reality, mixed reality can combine what you live with what you learn, effectively bridging the gap between how you learn something and how you employ it in the real world. Furthermore, given that mixed reality headsets often employ computer vision and tracking, these technologies have a heightened sense of the user and his/her body, affording the use of these systems to make use of the body as an interface to learn. This affordance can be combined with the immersive, highly-contextual nature of mixed reality platforms to create powerful systems that support learning inside and outside of the classroom. The purpose of this work is to explore mixed reality as a platform that is particularly suited for language learning, based on both its contextual and embodied capabilities, irrespective of its novelty as a platform.

In this thesis, our goal is two-fold. First, we aim to explore mixed reality as a learning platform with benefits that span beyond its widely explored engagement and contextual affordances, ratifying the platform as uniquely suited for embodied learning. We focus this exploration within the realm of language learning as an application space that is particularly fit for the affordances these technologies offer. This exploration is done through the development of two virtual reality systems that approach embodied learning in fundamentally different ways: explicitly using the body to encode new information and using the body to alter a student’s perception of self in order to change the way they engage with the learning material.

1.2 Motivation

Early childhood is full of experiences that employ the use of the body to learn [7, 8]. Children learn to use their fingers to deal with simple math operations. Or play games that involve motion to understand new concepts. This is also true for language learning in early stages of childhood. Children stories employ images, sounds, and tangible elements to
provide context, enhance memorability, and complement storytelling [9]. These elements are sometimes accompanied by enactment carried out by the readers and the listeners; helping children in early stages of literacy to highlight and cement different aspects of the material [10, 11].

Embodiment theories have shown that language is directly connected to the body [12], making the use of the body a promising venue to enhance language education. The connection between the body and language has been shown to hold for second languages [13], suggesting applications in this domain that could facilitate the way students learn. Despite the fact that there is opportunity to tap into the physicality of language to foster memorable educational activities, most language curriculums are primarily audio-visual in nature [4, 5], and have students memorizing vocabulary lists, and engaging in superficial exchanges in class that can quickly fade with time. After all, there are logistical limitations in the classroom that inhibit the effectiveness to which students can leverage their body to its full extent. Physical action is limited by space, and it is harder to guide a large group of students employing motion as a technique to encode new knowledge. Context is reduced to what is available in a room, limiting the opportunity of novel interactions that are relevant to the learning material.

Virtual reality systems can potentially bring back the body into the classroom and impact language education. Due to virtual reality’s immersive nature, it can allow a student to engage in activities that strongly relate to the vocabulary they are learning (e.g. learning vocabulary related to cooking while preparing a dish in a virtual kitchen). The tracking capabilities of virtual reality headsets can be leveraged to understand a student’s movements in space and provide real-time feedback, for example, generating second language descriptions of the user’s actions, thereby reducing the load on teachers during class. Students are no longer inhibited by their physical limitations, they can perform impossible actions that defy even gravity. Students can even change the way they look in VR to adapt the way they approach learning on a subconscious level.

Virtual reality systems have been employed to create embodiment illusions [14], where the perception of self is altered by changing the avatar a person inhabits in a virtual world.
This has been shown to alter the person's behavior, perception, and even their way of thinking or approaching problems [15, 16, 17]. Therefore, virtual reality does not only offer a way to enhance learning through the use of the body explicitly as a tool to encode new information, but it can affect the way a student behaves and approaches learning, using the body as a tool to subliminally enhance language acquisition.

1.3 Contributions

In this work, we discuss systems that take advantage of embodied elements to enhance the learning of a second language, and assess the effectiveness of embodied language learning in virtual reality with quantitative and qualitative data obtained from two user studies comprised of 60 participants each for a total of 120 subjects. We created two prototype systems to test two fundamentally different approaches to embodied language learning, i.e. using the body explicitly (Words in Motion) or implicitly (Inner Child) to enhance vocabulary acquisition. These systems can be used to teach students in a classroom or perform further experiments to understand how the body can be employed to teach more complex layers of language in more effective ways.

**Words in Motion** is a platform that presents a kinesthetic approach to enhancing the acquisition of vocabulary, using the body as a tool to explicitly encode new vocabulary and thereby enhance retention and recall. The system recognizes the learner's actions performed with virtual objects and presents the student with the target language word for a performed action (e.g. drink, pour, swing, etc.). Results from our experiment suggest that kinesthetic language learning in mixed reality enhances retention, with actions performed inside virtual reality increasing the recall rate of words a week after exposure.

**Inner Child** challenges the conventional approach of embodied learning, implicitly changing the way students learn by embodying them in the virtual avatar of a child. The system takes advantage of the embodiment illusion (i.e. inducing illusory ownership of a virtual avatar's body) [18] to alter a student's perception of self. Results from our user study suggest that simply having a virtual body impacts the acquisition of vocabulary in virtual reality in a positive way. Furthermore, the illusion of being younger in VR improved
participants' performance in subsequent recall tests both immediately and one week after exposure to the system, which was dependent on the type of body a participant inhabited in the virtual scenario. These results can radically change the way we learn in the classrooms of the future, suggesting that we can alter our physical notion of self to enhance the way we approach learning on a subconscious level. Because of the subliminal nature of this platform, it could be combined with explicit embodied methodologies in an immersive environment to create a powerful learning enhancement tool.

The systems presented in this work, along with the evaluations of these systems in both user studies, clearly show that virtual reality holds potential as a platform for embodied language learning. Furthermore, they suggest that the approaches presented in this thesis can be extrapolated to other fields of learning beyond language to impact education on a more general scale.

1.4 Thesis Overview

Chapter 2 begins by providing an overview of the connection between the mind and the body, focusing on how the body has been known to influence the way we learn and encode information. We provide a brief account of theories in the field of psychology that influenced our research and motivated the exploration of embodied applications for education in virtual reality. We present the relevant theory of embodied learning, and describe its applications within the realm of language learning. Explicit embodied and Implicit embodied learning are defined as fundamentally different approaches to embodied learning in virtual reality. We conclude this chapter by listing some of the unique traits of virtual reality that make it fit to engage students in embodied learning.

Chapter 3 presents Words in Motion, a prototype system that enables explicit embodied language learning in virtual reality. We detail the implementation of the system and include the design of a learning activity that incorporates it into a traditional classroom setting. We include the preliminary results from a pilot study with students who tried out the system and provided qualitative feedback about their experience.
Chapter 4 presents a formal evaluation of explicit embodied language learning in virtual reality. We provide the results and analysis of a study with 60 participants, testing the effects of explicit embodied learning with participants learning 20 new vocabulary words in Spanish using the *Words in Motion* platform.

Chapter 5 presents *Inner Child*, a virtual reality system that embodies users in the avatar of a child to allow implicit embodied language learning. We detail the design and implementation of the system.

Chapter 6 covers a formal evaluation of implicit embodied language learning in virtual reality. A study of the *Inner Child* experience, with 60 participants learning 20 vocabulary words in Spanish, is discussed in this section to highlight the effects of implicit embodied learning in virtual reality.
In this chapter, we will present research in the field of embodied cognition as well as its applications for learning. We will define embodied learning, which leverages research on embodied cognition to engage students in activities that involve their bodies to enhance learning. Furthermore, we dedicate a section to outline how embodied cognition and embodied learning have been applied to enhance language education, with supporting evidence that the body and language are connected. We define two distinct approaches to embodied learning which we address in the two prototypes built and the subsequent studies presented in this thesis. Finally, we present virtual reality as a platform which uniquely affords embodied language learning.
2.1 Embodied Cognition and Language

Embodied cognition is a branch of Cognitive Science research that proposes that human cognition is centered in the body [19]. That is, the way we acquire knowledge, think, and understand our experiences on a daily basis is leveraged by our whole body, and not only our brain [20, 21]. In biological terms, this can be explained as a consequence of the way a biological entity is able to interact and reason with its environment, which is inherently limited to its senses and physical constraints [22]. Nevertheless, supporting evidence suggests that even when an individual is decoupled or isolated from the environment, cognition is still grounded in mechanisms meant for interaction with the world [21].

Language has been shown to be linked to our bodily experiences [12], and some researchers include it as an extension of embodied cognition [23, 24, 25]. That is, the cognitive processes through which we ground language are directly affected by our physical reality and how it relates to our bodies. Even the way we structure our metaphors in language directly aligns with body-centric associations (e.g. “up” is good, “down” is bad) [26]. This connection is so intrinsic that brain imaging has shown that sensorimotor regions in the brain associated with carrying out a task light up when words associated with the corresponding actions are used or heard [27, 28, 29, 30].

![Figure 2](image)

Figure 2. The way metaphors are formulated directly relates to our physical body. In this case, known is represented as down, as it is physically reachable, while the unknown is conceptually linked to being up or physically out of reach [26].
A study found that emotion words (e.g. love) also triggered motor regions in the brain associated with common emotional behaviors performed with the arms, face, and hands, suggesting these effects hold even for more abstract words that are not directly linked to explicit action [31]. Moreover, studies support the notion that this relationship holds beyond our first language [13]. Learners presented with spatially suggestive words (e.g. “sky” strongly relates to the notion of “up”) show no difference in reaction times when asked to either raise or lower their hands in response to the perceived spatial association in two languages, suggesting no discrepancy between the bodily encoding of first and second languages [13]. The enactment effect showed that performing a set of actions can generally increase the capability of subjects to recall these tasks [32]. Given embodiment’s effects on memory, its impact on second language learning has also been studied within the context of vocabulary acquisition [28, 33, 34, 35]. Use of iconic gestures or illustrative motions have shown higher learning gain and retention than audiovisual modalities [36, 37, 38].

Different forms of language learning that involve the use of the body have found their ways into the conventional classroom. Total Physical Response (TPR) was introduced as a kinesthetic framework that consisted of teachers giving spoken commands to be performed by the student [39]. This created associations between the spoken order and the subsequent physical response of the learner. However, the framework was not formally evaluated and remained theoretical in nature [28, 39]. The use of dramatic enactment in second language learning has also been employed in the classrooms with positive effects in facilitating communication amongst students [40].
2.2 Embodied Learning Modalities

Embodied learning refers to learning that involves the body as a tool to enhance learning activities [41]. This branch of learning is informed by the theories that support embodied cognition, and often stems from constructivist models of education [42]. Like constructivist approaches, which involve students creating and experiencing to learn, embodied learning often uses bodily actions (e.g. making, building, enacting, etc.) as tools to enhance both the ways students approach learning and the learning outcome. In this section, we make a clear distinction between two different modalities through which embodied learning can be implemented in the classroom.

2.2.1 Explicit Embodied Learning

We categorize embodied learning as explicit if it uses the body as a tool to directly code new information, that is, learning that involves bodily action or form to create memorable recall cues that can be used to grasp new concepts. Kinesthetic learning, which we classify as explicit embodied learning focused on motion, can lead to increased comprehension and retention of the material [43, 44, 45]. Enacting is a well applied instance of kinesthetic learning, where students recreate a scenario or event in order to learn a new concept [40]. Total Physical Response [39] similarly encodes action verbs and phrases in a second language using a corresponding physical action.

Dual-coding theory proposes that information is embedded in both verbal and nonverbal representations within our brain [46]. These representations can be bridged by creating associations between the verbal component of a concept (e.g. the word “car”) and a non-verbal modality such as its visual analogue (e.g. an image of an old Volkswagen.) In theory, the more associations you create between both representations, the more likely you are to remember the concept in the long run. Existing mediums already exploit the power of strong visual associations to help learners achieve their goals (think of children’s books, which are full of pictures [9]). Kinesthetic learning’s effect on learning aligns with Dual Coding Theory; by performing an action and associating it with a new material, a new recall cue is established, grounded on the sensorimotor system.
Figure 4. Dual Coding Theory supports the notion that connections between verbal and non-verbal stimuli can be used to strengthen the memorability or recall of vocabulary words. Creating more referential connections between the verbal (word) and non-verbal (actions, pictures, sounds) systems theoretically leads to enhanced recall.
Not all learning that involves movement is necessarily considered kinesthetic and consequently in the realm of embodied learning. A student learning how to read, moves the pages in a book, and those engaging in learning activities that involve a computer oftentimes use their hands to move the mouse or interact with a touch screen. Although these learning instances involve the body and movement, they are not typically considered instances of kinesthetic or embodied learning. A key element in kinesthetic learning that differentiates it from other gestural movement is gestural congruency [47].

Gestural congruency refers to the relatedness of a performed action to the concept or material that is being learned through that action [47]. As we mentioned, a student flipping through the pages of a book involves action, but the action of flipping the page has nothing to do with the material he is reading. In order to achieve gestural congruency, the action that is being performed needs to hold a conceptual relationship with the material that is being learned. For instance, a teacher can encourage a student to clap for each letter or syllable of a word when they are practicing spelling. In this case, each clap is conceptually linked to a letter or syllable in the word, i.e. there is congruency between action and concept. Similarly, if students used their bodies to make letter shapes in order to learn how to write a character, there would be a conceptual relationship between body action and material. However, if this body action does not reflect the way the body is used when doing the task, there is no veridical convergence.

In this work, we extend the definition of kinesthetic learning to also include an element of veridical convergence. Veridical convergence refers to the relatedness of a performed action with the way the student employs it in real life. In other words, how the action that was used to learn a concept relates to how that concept is used in real life. In our previous example, where students learn to write characters by using their body to shape that character, there is gestural congruency, but no veridical convergence. When the students need to write down a character after they’ve used the aforementioned kinesthetic method, they don’t employ their body to shape the character.

A good example of veridical convergence is the way boxers practice fights by shadow boxing, performing actions such as jabbing and dodging the same way they are to be
employed when they execute the learned task. Total Physical Response is also a great example of kinesthetic learning with veridical convergence. Students learn new words by following an order in a second language which is directly linked to the words they are learning and how those words would be used in conversation to elicit the performed action. Oftentimes, kinesthetic learning that satisfies veridical convergence is by consequence gesturally congruent, but not the other way around.

Although the term kinesthetic learning is often used interchangeably with the term embodied learning, we consider kinesthetic learning as a subset of explicit embodied learning. Other explicit uses of the body that don’t involve action can be exploited to engage in embodied learning. For instance, the body can be used to represent concepts and create associations that don’t necessarily involve movement itself. The body can be used as a representational source, much like students use their fingers to learn how to add or subtract in early childhood, which isn’t necessarily associated with movement. Methods like the right-hand rule have found their way into the instruction of physics and mathematics, explicitly using the body to facilitate pedagogy of a complex concept by leveraging the spatial properties of the body rather than actual physical action.

![Figure 5. Examples of how the body form can be used as a representational medium to learn new concepts. (Left) The right-hand rule is used to understand the direction of the cross product between two vectors, with applications in physics and many other fields. (Right) Hand is often used by children as a method to help solve simple arithmetic problems. Image created by user:Acdx [GFDL (http://www.gnu.org/copyleft/fdl.html) or CC-BY-SA-3.0 (http://creativecommons.org/licenses/by-sa/3.0/)], via Wikimedia Commons.](image-url)
2.2.2 Implicit Embodied Learning

Most references in the literature regarding embodied learning focus on kinesthetic activities or explicit embodied learning instances [41, 48]. However, in this thesis we describe a distinct approach that can still be considered embodied learning even though it does not directly involve the explicit use of the body to learn, but rather its implicit or almost subliminal use. Slater presents the notion of leveraging illusory body ownership in virtual reality to change behaviors, and proposes the use of this modality to support implicit learning [49]. We define implicit embodied learning as a branch of embodied learning that uses the student’s perception of their body to change the way they engage with learning activities and the concepts learned. As opposed to explicit modalities, where the effects on retention and learning gains are primarily based on the student using their body to encode information, implicit learning does not require the student to actively reflect on the body or movement. Instead, perceptual changes of the body induce students to approach learning differently or with a different mindset.

Implicit embodied learning is therefore more nuanced, case-based, and consequently harder to employ in traditional curriculums. Without the use of technology, changing a student’s perceptions of their bodies can be challenging. However, implicit embodied learning occurs to some degree during activities such as enactment. An example would be students learning about a particular conflict in History class. Many times, by taking on the role of one of the sides, a student is able to better understand motivations behind a conflicting party. By creating the illusion of being someone else (e.g. through the use of costumes), a person can change the way they approach learning about their view of the world. This technique, often casually referred to as “walking in someone else’s shoes”, is an example of how we can change the body to enhance the way we learn implicitly. Similarly, depravation of senses such as sight, can result in a heightened sense of hearing [50], which can lead to people learning how to discern their environment differently, and even change learning of subjects like music through enhanced pitch discrimination [51]. This ties back to the notion of embodied cognition, which posits cognition is heavily influenced by the affordances and limitations of our physical body.
In this work, we categorize the effects of implicit embodied learning as either inward or outward facing. These effects can either be caused by factors directly related to the body and its affordances (inward facing), or to how society's view of the body influences the way we act (outwards facing).

Implicit embodied learning that is outwards facing in nature stems from a social aspect, concerning the value society places on the character our body represents. Slater embodied people in virtual reality as Albert Einstein, which in turn had participants performing better in cognitive tests [52]. This change in performance was associated with a boost in confidence and self-esteem. Clearly, embodying an avatar of a person with a gifted mind does not spark your body to perform better because of its inherent capabilities, but by the value society (and us as a result) give to the abilities of someone that looks like Albert Einstein. In other words, the learning enhancement comes from an external source, which in this case is society's view of Einstein's intelligence.

Implicit embodied learning that is inwards facing stems from inner ability of the body we own and is not related to social factors. The previous example about students learning about a character's motivation by enacting him or her is an example of implicit embodied learning that is inwards facing in nature. The enhanced ability to understand and learn about the character's motivation does not stem from an external societal influence (albeit it could under some circumstances), but rather due to our inherent ability to empathize more strongly with in-group members, or individuals we feel are more similar to us [53, 54]. Similarly, changing the affordances a virtual body has to engage in a learning activity in such a way that the participant grasps concepts differently, is related to the physical affordances of the body rather than an external source. For example, if the voice of a participant is modified such that its pitch sounded cartoonish, and this consequently leads to the participant speaking more in a language learning related task (perhaps because the voice is funny), the change in the way they approach learning is due to their physical affordances (in this case their voice). The enhanced learning occurs by factors that are inherent to our notion of self or our physical limitations, which are separated from an external influence of social norms or perceptions.
2.3 Virtual Reality Affordances for Embodied Learning

Virtual reality has been explored as a platform that supports embodied learning due to its immersive nature and tracking capabilities [48]. Applications have taken advantage of this to support learning of activities that are inherently physical in nature [55, 56, 57]. As a result, virtual reality has been used to support explicit embodied learning in different domains. In this section, we outline several features of virtual reality as a platform that we argue make it particularly suited to support explicit and implicit embodied learning.

2.3.1 Fine-Grained Body Tracking

Many virtual reality devices support body tracking. Although lower end platforms merely use a phone screen combined with optical lenses to provide a low-quality but cheap reproduction of virtual reality (e.g. Google Cardboard [58]), these devices are increasingly using computer vision to track a user's position in real space. Higher end devices, such as the HTC Vive [59], provide fine-grained body tracking which most commonly tracks a user's head and hands. Moreover, the release of additional trackers [60] and even body suits such as the Perception Neuron [61], enables multiple point tracking of the user's body.

Sensing modalities that have become more popularized in the past year include other bodily signals that span beyond spatial understanding of the user's body, such as electrodermal activity, electroencephalography (EEG), electromyography (EMG), and eye tracking. Because virtual reality uses a head mounted display, users are more traditionally disposed to placing the device on their face, an opportunity that researchers have used to embed all these sensors packaged with the HMD, taking advantage of its direct contact with the skin [62]. This allows virtual reality applications to be seamlessly hyper aware of the user's body and cognitive state, unlike other platforms such as traditional desktop and mobile modalities. As a result, virtual reality applications can support embodied learning by enabling the development of applications with real-time information about the user's position, movement, and cognitive state.
2.3.2 Contextual Affinity

Virtual Reality can be employed to take users on a journey to distant lands, immersing student’s in otherwise unreachable places from the comfort of the classroom or home. This, in itself contributes to learning of sorts that is not necessarily embodied in nature (e.g. Contextual/Situated Learning [63]). However, the ability to place students in locations contextually relevant to the material they are learning, can supplement explicit embodied learning by enhancing veridical convergence. Participants performing actions to learn a concept can strengthen the action within a context where the concept is executed. For instance, learning vocabulary related to cooking in a virtual kitchen, as opposed to in a traditional classroom environment. Moreover, context can also be employed as a powerful tool to alter a user’s sense of self [64], where subjects can use their virtual identity to define who they are in a particular scenario. For instance, being an adult in a virtual classroom full of kids can support the illusion of being a teacher, while being a child in said environment can potentially enhance a sense of youth. Like we previously mentioned, changing a user’s perception of self can indeed affect the way they engage with the world and consequently with learning tasks [49]. Therefore, virtual reality’s ability to seamlessly immerse student’s in new locales or situations can impact both implicit and explicit embodied learning.

2.3.3 Enhanced Action\Representation Space

Embodied cognition supports the notion that the way we interpret the world is tied to our physical limitations [19]. As a consequence, the extent to which we can use our body or its representation is limited by the constraints of our body. Virtual simulations are not limited by constraints such as gravity. A virtual avatar’s speed or flexibility is not hindered by the user’s stamina or physical constitution. Actions that could not be otherwise performed by a student in the real world, can be executed with ease in virtual reality. If actions can be thought of as an encoding space, then incrementing this space allows for more possible encodings for new concepts. Not to mention that the memorability of “impossible” actions could be taken into advantage to create more meaningful and memorable encodings.
Even further than increasing the space of actions that can be employed for explicit embodied learning, virtual reality allows an extended representation of the body. This can enhance both explicit and implicit embodied learning activities. A user's virtual avatar can be changed to a form that is not human, or that extends the human body. Research has found that the mind can adapt to incorporate body representations that don't necessarily converge with our own. For instance, participants in virtual reality have quickly adapted to performing tasks more efficiently when their virtual self has an additional limb [65]. This allows an extended representational space that can be harnessed for embodied learning. Take for instance the way children learn how to add and subtract using the fingers in their hands. This technique is limited by the amount of fingers a child has. If an extended representation of the hand in virtual reality were to include additional fingers, these could enhance the technique. Similarly, an altered representation of the body can potentially change the way we approach learning tasks and implicitly enhance embodied learning.

2.3.4 Becoming Others

The Proteus effect describes how users in virtual spaces tend to behave in a way that adapts to the characteristics of their avatars [66]. This change in behavior is associated with the ratification of stereotypes, by which a user behaves the way he/she believes a character with certain traits is expected to behave by others, which is explained by Self Perception Theory [67]. For instance, participants who inhabit a virtual body considered attractive interact more confidently and extroverted in social scenarios [66]. The adoption of avatar traits can be beneficial in certain contexts. Participants who embodied a taller avatar in a study, were more aggressive negotiators when bartering, which can result in a more favorable outcome for the subject [66]. Taking advantage of the Proteus effect to enhance learning is an example of implicit embodied learning that is outwards facing in nature. Because virtual reality offers full body immersion/tracking, it enables users to inhabit anybody, human or not, and potentially adopt behaviors related to the inhabited body that can be leveraged for implicit embodied learning.
Chapter 3
WORDS IN MOTION

Figure 6. In the Words in Motion platform, actions performed in the virtual environment are recognized to reinforce vocabulary learning of verb-action pairs.

Technology enabled embodied learning has been explored as a means to enhance learning within different disciplines. The most common usage is learning tasks which are characteristically physical, such as meditation [56] or sports [57], due to the capability of tracking technology to provide real-time feedback for pose-related activities. Nevertheless, tracking systems like Microsoft’s Kinect have been used to teach in domains that are not characteristically physical, leveraging embodied cognition to teach chemistry [68], math [69], and anatomy [55], among others. In this chapter, we focus specifically on the development of a novel explicit embodied language learning platform named Words in Motion. Words in Motion recognizes actions in the virtual space and presents students with the second language word associated with it to enhance vocabulary acquisition. We describe related work in the space of explicit embodied language learning in virtual reality, the design and implementation of the different components of the platform, and a pilot study of a learning activity supported by Words in Motion.

3.1 Related Work
A large body of work has been dedicated to learning applications that involve explicit embodied elements to enhance language education. To guide our discussion, we divide
these into three categories based on the relationship between the embodied elements and the learning material.

3.1.1 No Gestural Congruency, No Veridical Convergence

Systems in this category use physical action that is not related to the learning content and can be classified more as gestural interaction with the material. Ogma [70] is a virtual reality system that uses the MYO armband to allow students to navigate a virtual environment and learn new vocabulary words that are related to that environment. Similarly, PILE [71] creates mixed reality environments that can be manipulated with physical gestures to engage in different aspects of language learning (e.g. moving hands to match pictures with words in the target language). JaJan! [72] takes a similar approach, but involves the presence of a co-located or remote peer to enhance the learning activities.

3.1.2 Gestural Congruency, No Veridical Convergence

This category covers projects that leverage action that is linked to language but is not strongly associated with how the linguistic component is employed in the real world. I.e., there is a relationship between action and words (gestural congruency), but this relationship is loosely coupled to how the action is employed in the real world (low veridical convergence). An example of this is WordOut [73], a project turned museum exhibit in which children can use their body to match letter shapes as a literacy enhancing activity. Cheng et al. [74] presented a virtual reality game that incorporated body movement related to cultural components of language. The system primarily focused on bowing during social interactions with virtual avatars to enhance the learning of language and culture.

Although evidence suggests that experiences in the first two categories also benefit from kinesthetic encoding of knowledge, the explicit embodied learning experiences presented in this thesis are concerned with technology enhanced kinesthetic activities that create direct associations between actions and words.
3.1.2 Gestural Congruency and Veridical Convergence

This category covers projects that teach language with a direct relationship between bodily action and meaning, with coherence between employed action and real-world usage. *SpatialEase* was a Kinect game developed by Microsoft Research where the participants respond to audio cues in Mandarin with bodily actions [75]. *SpatialEase* was compared with Rosetta Stone software, showing similar gains for vocabulary and grammar acquisition. Similarly, Kuo et al. [43] developed a Kinect based platform to compare the effectiveness of computer enabled Total Physical Response (TPR) [39] versus the classical TPR method. Testing with 50 elementary students showed no significant differences between conditions immediately after the short learning session. Nevertheless, the retention of vocabulary was shown as increased for the technology enhanced group. Similar approaches, making use of Kinect or alternate tracking systems, are presented in [44, 45].

Words in Motion enables embodied language learning that exploits a direct relation between body action and language. Our platform leverages kinesthetic experiences in virtual reality, whereas existing platforms [43, 44, 45] emphasize non-immersive analogues. Comparisons between computer enabled and classical TPR suggest that there is a difference in the learning gain, despite both activities requiring the same bodily motions [43]. This implies that the modality through which the kinesthetic learning occurs impacts the way material is encoded in bodily motion, and motivates an exploration of virtual reality as a platform for kinesthetic language learning. Although virtual reality is fundamentally an embodied platform, to the best of our knowledge, it has not been used for language learning that involves the use of the body as a tool to explicitly encode related vocabulary, which is the focus of the prototype presented in this chapter.
3.2 System Implementation

*Words in Motion* is a system developed using the HTC Vive that enables kinesthetic language learning in VR. The platform was developed as a recognition system that allows teachers to engage students in activities that introduce kinesthetic elements to second language instruction.

3.2.1 Recognizing Actions

*Words in Motion* augments objects in a virtual environment with verb-action pairs to reinforce associations between words and the body. Learners can perform actions within the virtual environment that trigger the corresponding name of the action to appear briefly before the student, thereby using the body as a tool to encode new vocabulary. The recognition software is a heavily modified version of Edwon’s ultimate gesture VR plugin [76]. Every object in the virtual environment is supported by a feed-forward neural net trained to recognize actions that can be performed with it.

Users can begin the recognition process by pressing the Vive controller's trigger on the hand holding the virtual object. This provides haptic (a subtle vibration) and visual (a transparent trail) feedback to denote an action is being recorded. When the trigger is released, the path is evaluated and classified. If the path matches a trained action-word pair with a degree of confidence, the corresponding word appears floating momentarily in front.
of the user (Figure 7). Currently, the system is limited to motions with the user's hands. For instance, users can perform motions with a virtual brush to learn the Spanish word “pintar” (to paint). Minimal changes would allow the use of Vive trackers to recognize actions that involve full body motion, such as kicking a ball or dancing.

3.2.2 Training New Actions

Teachers can enter the virtual environment to create new actions for students. In order to augment an object with new word-action pairs, a teacher can grab it from the environment and perform the new action multiple times to “teach” the neural net by example to recognize this action. The teacher can perform the action a variable number of times, but generally 10-15 examples is enough to create an accurate model. Training a new verb-action pair for the object, creates a motion signifier. This signifier is a dynamically generated animation that displays the characteristic path that would trigger the word to appear. An orb travels along this path to show students how to perform the new action.

3.2.2 Challenges

The system supports the creation of action lists or “challenges” for students. A challenge consists of a sequence of actions that must be executed by a user with a specific set of objects in the virtual context. A desktop interface allows teachers or students to create these challenges to engage in game-like activities. These activities can be contextualized to combine situated and kinesthetic learning and further enhance the language learning experience. The challenge interface automatically populates a list of all the objects in the virtual scene that are actionable by the students. The user can select an object from this list, which automatically populates a second list of actions that can be performed with that particular object. Students can then drag the action they want into a queue that determines the action sequences that the participant must perform. When the user is finished selecting the sequence, the participant inside virtual reality is informed of the actions he/she must perform in order to successfully complete the challenge.
Figure 8. A desktop interface allows teachers or students to create new action-word pairs for other participants.

Figure 9. The designed activity takes place in a virtual kitchen, where the student must prepare a plate to complete the challenge.
3.3 Learning Activity Design

We designed an activity that leverages the kinesthetic capabilities of *Words in Motion* but also addresses some of the disadvantages noted in purely kinesthetic experiences like Total Physical Response [39]: the lack of conversational interactions. The motivation was to engage learners in a game that incorporates both kinesthetic and conversational elements in a way that could be employed as a classroom activity that would be deployed to Kanda University in Tokyo to support English instruction of college level students. A virtual kitchen environment was created and a set of objects were pre-trained with contextual action-word pairs (e.g. chopping with a knife) (Figure 10). The game's goal was to perform a sequence of actions in the virtual kitchen environment with the correct set of objects.

The activity was designed with multiple players in mind, where one player is immersed in VR, while the others participate from the real world. This was purposely done to address the fact that many educational facilities have constraints that limit the number of room-scale devices they can afford or accommodate in a classroom. The participant inside the virtual environment takes the role of the "performer", while the participant(s) outside VR are denoted the instructor(s).

The role of the "performer" is to execute the right sequence of actions using the target objects in the kitchen (Figure 10). However, there is no indication within the virtual environment that informs the performer (i) which action needs to be performed, (2) how to perform the action in space, or (3) which object to perform the action with.

Participants outside the virtual environment take on the role of the instructors. Instructors have two views on external monitors. On one screen, they can monitor the
performer’s point of view. The other screen displays the motion signifier, an animation that shows the path that the performer must enact, along with an instruction of which object to perform it with. The role of the instructors is to communicate verbally with the performer the actions they have to perform, how to perform them, and which object to look for in the virtual kitchen.

There is no method that enforces how communication between the instructor(s) and the performer happens. This allows the teacher or moderator to set constraints that give the right measure of difficulty according to the participants’ fluency in the target language. This can range from full second language communication, to first language instruction (where only the performer learns the target language by kinesthetic means). This allows teachers to engage the whole classroom as instructors that practice conversationally by communicating in the second language with the performers, while taking turns to engage in kinesthetic reinforcement of the material using virtual reality.

### 3.4 Activity Trials

The *Words in Motion* system was tested with 11 students under varying conditions in a pilot study. Out of the 11 participants, four acted only as performers while being instructed by a researcher to perform actions. Remaining subjects participated in the activity in pairs, taking either the role of the performer or of the instructor.

Participants were to finish a game level that consisted of 5 unique actions, in a set of 6 total actions (one of the actions is required twice.) All of the actions were contextually relevant to the object they had to be carried out with (e.g. chop with a knife). Users with enough proficiency were asked to communicate fully in a second language, but were allowed to revert to English if they became stuck or were not able to transmit the motion path effectively with their limited vocabulary. Participants were interviewed briefly after the experience to obtain feedback on their experience.

All of the performers agreed that the controls were intuitive. Users required little to no training, in order for them to fulfill all the required actions. One of the trials consisted of a participant that had already acted out the role of a performer. As an instructor, she was
able to explain actions much more efficiently to the other player. This suggests she was able to create familiarity with the system in a relatively short amount of time, such that she was able to transmit that knowledge much more effectively than instructors who had never acted out as performers.

Although the participants described the controls as intuitive, a common theme emerged from conversations with the immersed subjects. Out of the 8 performers, 6 noted that the way that actions were requested to be performed did not match their mental model of how that action was performed in the real world. For instance, “chopping with a knife” was expected by the participants to be short arc motions from left to right. The system, however, required users to move the virtual knife up and down. One subject mentioned that giving an action a name made it harder for him to perform the action if it didn’t converge with his mental model of it. Another participant expressed the realization that the game was, “not about performing actions as I envision them, but actions as the instructor wants them.” This made it harder for users to perform the actions, which in turn detracted from the vocabulary association and focused attention primarily on the action. This posits an interesting question regarding the effectiveness of the kinesthetic approach that is recognized automatically, as opposed to qualitatively judged by a teacher: Is kinesthetic learning effective, despite diverging models of action across students?

Appearance of the second language word and the audiovisual cues given by the system when the correct action is performed, played an important role for the students. In the experience, a success sound was played every time the correct action was performed. The corresponding word in the second language also appeared in front of the user, momentarily floating away from the hand with which the action was performed. Participants report ignoring the sound and relying on the word for confirmation that the correct action was performed. However, several participants expressed the desire to see the word before they performed the action as opposed to briefly exposing it right after.

In the following chapter, we show the results of a user study in which we formally evaluate the effects of explicit embodied language learning using the Words in Motion platform to teach students 20 new vocabulary words in Spanish.
Chapter 4
EXPLICIT EMBODIED LEARNING STUDY

4.1 Vocabulary Acquisition Study Design

Words in Motion was tested during a controlled experiment with 60 participants subjected to a vocabulary learning task. The task consisted of learning a set of 20 new vocabulary words in Spanish under three different conditions: outside of VR, in VR without performing actions, and in VR using the Words in Motion platform. Subjects in all conditions were subsequently evaluated to determine the effects of the Words in Motion platform on vocabulary recall tests. The following section describes the rationale behind the word selection and evaluation methods for the controlled experiment.

4.1.1 Word Selection

Participants were exposed to a set of 20 transitive verbs in Spanish. The words were selected from Lifcach [24], a frequency list with over 450 million Spanish words. We selected words in low frequency bins in order to ensure the task would be challenging even to participants with prior exposure to Spanish. An action was then chosen in the Words in Motion platform for each word and the neural net trained to recognize that action gesture. This action directly matched the meaning of the paired word. We removed any English cognates from our selection to reduce the influence of prior knowledge on the results.

4.1.2 Evaluation Method

The students were administered tests that required them to provide the English translation of a word after being prompted with the Spanish analogue. The tests consisted of 25 words (which included words that the participant would not learn throughout the experience as a mechanism to ensure subjects were not cheating in remote delayed tests). Participants were instructed to leave an answer blank if they did not know it.
4.1.3 Subject Selection

A study was conducted with 57 participants to assess the effectiveness of the kinesthetic method in virtual reality in comparison to a group which learned in a text-only condition outside of VR and a group that learned in VR with no kinesthetic component. Both text-only and VR kinesthetic groups consisted of 20 participants who were recruited from the university campus. The non-kinesthetic VR group consisted of 17 participants recruited from campus, as 3 participants dropped out before the final assessment. All participants were awarded a $10 Amazon gift card upon completion of the experimental procedure.

4.1.4 Experimental Procedure

Subjects were randomly assigned to one of the three learning conditions: a text-only condition, a virtual non-kinesthetic condition, and the virtual kinesthetic condition supported by our system.

The text-only group, would sit and try to learn the words by observing the set of 20 word pairs that cycled on a computer monitor. Each word pair consisted of the target word in
Spanish and its corresponding English translation. Each word pair was presented for 15 seconds, with each pair shown twice for a total of 30 seconds worth of exposure to each new vocabulary word.

The kinesthetic learners in virtual reality would be standing in an empty (virtual) room. The environment was kept minimal to mitigate the effects of context on the experimental results. Participants were instructed to perform the movement depicted by a motion signifier at least twice. The first movement triggered the associated word pair to appear in front of the participant. The second movement was requested to create a direct association between the word pair and the body action (since the target word is not visible during the execution of the first action). The word pair remained visible for 15 seconds, before the next action was queued. Each action was requested twice throughout the learning session, for a total of 30 seconds worth of exposure to each new vocabulary word. The non-kinesthetic learners in VR would go through the same experience as the VR kinesthetic learners, but instead of performing the actions themselves, they would see the virtual object performing the associated motion on its own. All groups were tested for knowledge of the words both immediately after the training session and exactly one week after, using the same test that was administered before exposure to the experimental condition.

![Word Retention](image)

*Figure 18. Words recalled by participants immediately after exposure and one week after.*
Figure 19. Percentage of words lost between immediate and delayed post-test show higher retention for VR Kinesthetic learners.

4.2 Results

Participants from all groups scored low on the pre-test, with the mean number of words identified in the pre-test for the text-only (M=0.65, SE=0.37), virtual non-kinesthetic (M=0, SE=0), and the virtual kinesthetic (M=0.25, SE=0.12) approach showing that very few participants had prior knowledge of the selected words. The mean total number of words recalled during the immediate post-test differed between the text-only condition (M=14.6, SE=1.15), virtual non-kinesthetic (M=9.41, SE=1.08), and the virtual kinesthetic condition (M=10.8, SE=1.23). A t-test revealed differences between group means, with participants in the text-only condition significantly outperforming non-kinesthetic (p=0.002) and kinesthetic learners (p=0.03) in VR. However, one week later, the mean number of words still remembered by the kinesthetic learners (M=7.8, SE=1.19) and participants in the control condition (M=7.56, SE=1.27) were virtually the same (see Figure 18). Virtual non-kinesthetic learners (M=3.18, SE=0.75) performed significantly lower than text-only (p=0.008) and virtual kinesthetic (p=0.002) groups a week after.

Figure 19 shows that the mean percentage of words lost between the immediate and delayed post-tests between text-only (M=0.54 SE=0.06), VR non-kinesthetic (M=0.59, SE=0.08), and VR kinesthetic (M=0.28, SE=0.08) participants showed that those involved in kinesthetic training retained more words on average than those in the text-only (p=0.006) and non-kinesthetic virtual conditions (p=0.01). No significant difference was found between the retention rate of text-only and non-kinesthetic virtual conditions (p = 0.79).
Correlation analysis was performed on the metadata obtained from the VR kinesthetic group. Due to minor issues with the telemetry collecting module, only metadata from 14 participants was analyzed as the data for 6 participants was partial or corrupted. This data included telemetry of the successful and failed attempts at performing an action. A moderate positive correlation was found between the amount of successfully performed actions and the times that word is remembered in both the immediate post-test and a week after (Figure 20). This correlation is stronger a week later ($r=0.67$) than immediately after the training session ($r=0.52$) ($p<0.05$).

Participants in the virtual reality condition completed a survey specifically addressing self-reported engagement (see *Engagement* in Appendix B) during the experience and perceived intuitiveness of the system. Users of the *Words in Motion* platform reported moderate to high engagement on a 5-point Likert scale. Likewise, the system's intuitiveness was perceived moderate to high amongst participants in the virtual kinesthetic condition.

Open commentary from participants in both groups was collected regarding their experiences throughout the experimental procedure. In the Virtual Reality settings, several users pointed out that auditory feedback or hearing the word would have made the task easier for them. One subject pointed out that he was not very convinced by the experience, but was surprised a week after when he was able to remember so many words. Some participants did not understand what the actions were for, pointing out that they felt like "something else to memorize". They felt actions were distracting them from learning the words. Particularly, participants pointed out being confused or frustrated.
when the action they had to perform did not match their mental model of how that action would normally be executed by them. This separation or distance was described by subjects as distracting and causing their focus to shift towards the action instead of the words to be learned. Nevertheless, positive feedback related to the actions was also observed among subjects, where they pointed out “visualizing themselves performing the actions” to remember the words. Participants in the text-only flashcard condition report the experience as not very engaging or memorable. Many felt it more like a cramming or memorization task, as opposed to a learning task given the lack of context.

### 4.3 Analysis

Immediately after they are exposed to the learning experience, participants in the control case significantly outperformed those who learned the words in both VR conditions (Figure 18). This result differs from prior literature, where kinesthetic learners often match or outperform non-kinesthetic methods [43, 45]. These results suggest that participants in VR are initially hindered or distracted by the novelty of the system. Given that both kinesthetic and non-kinesthetic VR groups are outperformed by the text-only group, we can assume that the kinesthetic component is not the confounding element. The fact that the number of failed attempts at performing an action has virtually no correlation (and more importantly no negative correlation) with the times the word is remembered in subsequent tests, further supports this statement.
A study in a VR platform called Ogma [70] that explored vocabulary acquisition among participants in VR versus subjects using flashcards, revealed results that are similar to the ones presented in our work. There are two main differences between that study and ours. First, we focused on second language acquisition using Spanish as the target language, whereas the Ogma study taught participants in Swedish as a second language. Second, participants in the Ogma study performed no actions to learn the vocabulary words; subjects simply navigated a virtual environment populated with labeled objects, using a MYO armband to control how they explored a virtual house. Authors of the Ogma platform also attribute the smaller number of words immediately recalled by students in VR to the familiarity students have with flash-card style learning.

Despite the initial advantage that the text-only method has over the VR kinesthetic approach, one week later there is no significant difference between the performance of both groups. However, there is a significant difference in the amount of words recalled between participants who engaged in kinesthetic learning in VR and those who did not (Figure 18). Correlation analysis between the number of correct actions performed for a word and the amount of times it was remembered both immediately and one week after suggests a relevant effect from the kinesthetic nature of the learning experience (Figure 20). Namely, the more times a subject performed an action in VR, the more likely that subject will remember the associated word. This correlation is stronger after a week, supporting the notion that the kinesthetic component plays a strong role in the retention of vocabulary, and explains why participants performing actions in VR remember more words than those in non-kinesthetic interactions. This aligns results from previous experiments that compare kinesthetic and non-kinesthetic methods [43, 44, 45].

Although VR kinesthetic subjects performed similarly to the text-only group a week after, the rate at which participants forgot the words they had learned was significantly lower for VR kinesthetic learners while it was similar between participants in the text-only and VR non-kinesthetic conditions (Figure 19). In other words, VR kinesthetic learning experiences were more memorable and helped participants retain a larger number of words, despite any confounding elements that hindered their initial learning gain.
Chapter 5
INNER CHILD

Figure 11. (Left) Mixamo auto rigging creates a bone structure from our designed models, allowing us to animate a body to the movement of a user.

Inner Child is a Virtual Reality experience developed on the HTC Vive that embodies users in the virtual avatar of a child. The platform leverages the tracking capabilities of room-scale virtual reality to create the illusion of feeling younger, so that learners can approach language learning with a more childlike mindset. Our hypothesis is that having a virtual body can impact the acquisition of vocabulary. We hypothesize that virtual reality can be used to decrease a user’s subjective age, by embodying them in the virtual avatar of a child, which will result in enhanced recall in vocabulary learning tasks. In this chapter, we cover the supporting evidence that suggests why embodiment as a child can enhance the way we learn, the implementation of the Inner Child system, and the rationale behind the design decisions for the development of a virtual classroom to create the illusion of being a child.

5.1 Becoming a Child (Again)
Children can be better at learning a second language than adults. Socio psychological factors often inhibit adults in learning environments [77, 78]; they are afraid to make mistakes or look silly around their peers and as a result don’t engage with the material as a
child would. In language learning, this becomes even more prominent, as children are less afraid of making weird sounds or mispronouncing new words, whereas older students hesitate to speak out loud [79]. Not only are children less constrained by socio psychological pressures, but they tend to be naturally more curious [80] and open to new ideas when learning [81]. They approach the task of learning, in many ways, different than an adult would [82]. The Inner Child platform intends to tap into the potential benefits of changing the way we perceive ourselves by altering our body and environment in virtual reality. The platform was motivated by two fundamental assumptions. First, the way we perceive ourselves affects the way we engage with the world. Second, the way we perceive the environment affects how we perceive ourselves.

An experiment that embodied subjects into the body of a child in virtual reality discovered that it affected the way subjects perceived the world [15]. Participants who were embodied as children overestimated the size of objects after being exposed to the embodiment illusion [15]. These results are congruent with [83] in which adults embodied as small dolls tended to overestimate the size of objects. However, the participants who were embodied in the avatar of children, overestimated the size of objects they saw in a virtual space by twice as much as subjects embodied in child-sized adult avatars. This suggests that the effect of overestimation is not only caused by a reduction in size of the avatar with respect to the environment, but by the shape of the avatar and its representation [15]. The illusion also affected the subjects embodied as children in Implicit Association Tests (IAT) [84], in which they self-identified with youthful terms on a subconscious level after being embodied as children [85]. A subsequent experiment replicated these results and showed that by also altering the participant’s voice to be higher pitch while embodied as a child, temporarily changed the pitch of their voice after the exposure [85].

Embodied cognition posits that the environment also plays a fundamental role in cognition [21]. Margaret Wilson establishes the environment as part of a cohesive unit with the body, in her six views of embodied cognition [21]. Therefore, it is not only the property of having a child-like body which can change the way we approach tasks, but also being in
the right environment can further enhance experiences. In Counterclockwise, Ellen Langer, a social psychologist, writes about an experiment conducted in 1979 in which she placed elderly men in an environment that recreated the appearance of 1959 (20 years earlier) [64]. Langer reports that the participants not only displayed benefits in health (better mobility, dexterity, lower blood pressure, etc.), but also “appeared” younger to others after this exposure [64]. These results suggest that we can somehow trick our minds to think differently based on who and where we are. But can we actually tap into our inner child and change the way we behave in learning tasks?

A study showed that just priming subjects with the idea that they are children can impact their performance when carrying out creative tasks [86]. Subjects taking a Torrance Creativity Test [87] who put themselves in the mindset of a child performed significantly better than adults who were not primed to think as if they were children when performing the test [86]. This was the case particularly for subjects who were introverts, suggesting that the effectiveness of this is dependent on the subjects’ personality [86]. Becoming a child in virtual reality can potentially be used as an implicit embodied technique to enhance language acquisition, which is the goal of Inner Child.

The literature supports the notion that changing a person’s perception of self is possible by changing the way we perceive our body, and by altering our environment [14]. Not only can this change the way we behave and perceive our environment, but it can also be used as a tool to enhance our performance in different tasks. The effect of embodying different avatars on cognitive abilities and self-perception has been widely explored [49, 85, 14, 52], but the applications of these effects on learning are underexplored. To the best of our knowledge, these techniques have not been applied in the realm of language learning.

5.2 Designs to Trick the Mind

Creating the illusion of becoming a child in the virtual reality experience is pivotal to exploring its potential effects on the learning task. For the mind to accept ownership of the virtual body of a child is a different challenge than to enter a child-like or more juvenile mindset. In the following section, we detail the interplay between the different elements in
the *Inner Child* embodiment illusion experience, and how they play a role in establishing the desired decrease in subjective age.

5.2.1 Visual-Motor Synchrony

The primary method for establishing the embodiment illusion in the Inner Child experience is visual-motor synchrony. By aligning the user’s movements with those of the avatar, and providing a feedback loop so that the user is able to acknowledge that he/she is in control of the body, we cement the illusion that in virtual reality this is indeed their body. Alternative methods have been explored to establish a strong sense of embodiment in a proxy including somatosensory stimulation using the sense of touch and congruent visuoproprioceptive cues [18]. The rubber hand illusion [88] is a good example of this approach, which conveys an effective illusion of ownership and has been reproduced in both real [89] and virtual settings [90]. Several user studies have shown, however, that visual-motor synchrony with a humanoid body in first person perspective is sufficient to establish the embodiment illusion in virtual reality [15, 52].

Inner Child uses mirrors to establish the body ownership illusion. Although users are able to look down at their bodies, the use of mirrors allows for a full body view, which in turn can create a stronger sense of presence. Although future iterations of the system could allow users to select their appearance to match their physical self, our design does not include appearance matching, under the assumption that the body ownership illusion has been found to be effective regardless of how similar (or dissimilar) your body proxy is to your real body. This has been explored for both real world illusions (e.g. rubber hand experiment with a hand that has a different skin tone [91]) and virtual experiences [15, 85] (e.g. avatar of children who look remarkably different than the subject).

5.2.3 Environment

The Counterclockwise experiments [64] highlighted the effect of the environment that plays into embodied cognition. Although not widely accepted by proponents of this theory, as Margaret Wilson points out when outlining the six views of embodied cognition,
cognition is offloaded on the environment and consequently the mind and the body are affected by situation and surroundings [21]. This is also supported by extended cognition, which more liberally attributes cognition to elements outside the mind and body [92]. Given that the Counterclockwise experiments showed that a highly congruent environment has the ability to affect a person’s perception of their age, we also decided to use the virtual environment itself to enhance the sense of being an adult or a child. For the Inner Child experience, we selected an elementary school classroom populated by a teacher and students. Additionally, sounds of chatter obtained from an elementary classroom is looped in the background to create a more realistic depiction of a school setting.

Figure 2. Inner Child uses a virtual classroom environment to help induce the illusion of being a child.

5.2.4 Voice

Although the use of different sensorial methods to establish the body ownership illusion have not been found as pivotal as motor visual synchrony, we believe that voice is pivotal when body ownership is used to enhance language learning. Experiments have found that, although changing the voice does not significantly affect embodiment in the avatar of a
child, a discrepancy between the voice and the body you own in virtual reality can detrimentally affect it [85]. Because spoken language has a prominent role in the classroom, not having the auditory feedback or voice of a child can hinder the user's ability to remain under the illusion throughout a learning session. Consequently, we used voice changing software to alter the pitch and formant of the user's voice in order to make participant's sound like a child version of themselves.

5.2.5 Scaling
One of the biggest challenges when creating the child embodiment illusion was scaling. Visual-motor synchrony is enough to establish the body ownership illusion [16]. However, in other previous experiments where embodiment has been done in the body of a child, the illusion has been less powerful than that of embodying an adult [15, 85]. It does not come as a surprise that it is more challenging to create the illusion of being a child. After all, the discrepancy between a child and adult body is significantly greater than embodying an adult that looks different from the user, for example in skin color or gender.

Scaling is a two-part process in the Inner Child experience. First, the system needs to scale the avatar to the user's size (making it smaller or bigger depending on the situation). Subsequently, in order to keep the original relationship between the size of the avatar the user embodies and the environment, the environment must be scaled as well. In the Inner Child experience, the resulting effect is the feeling that the user has shrunk to the size of a child, when in effect, we have made the child body as large as the user's body and subsequently modified the environment to retain the same relative proportions. This mechanism is beneficial in two ways. First, the designer of the experience can be completely agnostic of the user's size in the design and implementation of a new scene. In essence, what he/she sees in the Unity editor will be kept true to the original intent of the developer. Second, it allows us to keep the kinematics used to actuate the avatar body unaffected by the dimensional changes.

We hypothesized that in cases where the body's dimensions diverge too much from those of the user, we would need to strengthen the experience by providing more cues that
reinforce the illusion. As a result, throughout the experience we created visual cues to ground the relationship between the user’s size and the world’s. Given one of the most notable differences (in physical terms) between being a child or an adult is height, we focused on elements that augment the feeling of being of a shorter size. This includes making objects around the user just slightly bigger than they would usually be. Since we make the avatar body as big as the user’s, without the proper cues, it’s easy to forgo the illusion that the user is indeed smaller. In the case of Inner Child, we strategically positioned objects in the environment, such as drawers, lockers, and even other human avatars that almost subliminally evoke a comparison between the user and the world.

5.3 Implementation

5.3.1 Platform

Inner child was developed on the HTC Vive, for its capability to provide fine-grained, room-scale tracking and high resolution virtual reality experiences. The Lighthouse room-scale tracking allowed us to create a virtual space where the user can move freely and interact with the environment without inducing motion, which in turn let us preemptively address motion sickness. The experience is not computationally expensive, running on a Mac Book Pro 2017 using boot camp Windows with 16GB RAM and a Radeon Pro 555 with 2GB of GDDR5 memory (which barely meets the recommended graphic standards for the HTC Vive platform). The system was developed in Unity using C# for scripting. Inverse kinematics were provided by Final IK [93].

5.3.2 Avatar Design

The compatible models and all characters used in scenes to enhance the embodiment illusion were designed in Adobe Fuse. Fuse is a character creator that seamlessly allows the creation of humanoid models. Although this software enables users to click and pick the appearances of a character (including clothes and accessories), it does not currently let users create child avatars. Therefore, a humanoid adult character had to be thoroughly edited to create the appearance of a child, by reducing the dimensions of the body and
softening its facial features. Female and male versions of the child avatar were created and subsequently rigged for animation. An adult version of these avatars was also designed to use for the control condition in following experiments. Six additional avatars (4 child and 2 adult) were also created to inhabit the environment and further the illusion of a real classroom. The features of the avatars are not editable by the user, given the literature supports the embodiment illusion can be successfully executed with bodies that do not match the subject's features.

Adobe Fuse allowed us to automatically rig characters in order to animate them using Mixamo’s web service. Mixamo is a web-based service that allows relatively accessible rigging and animation of humanoid avatars. The pipeline would be to create the character in Fuse, which automatically rigs and delivers the character to the Mixamo web service. Afterwards, the Mixamo platform let us download a rigged character that can be directly imported in a Unity scene. Problems with importing the Mixamo characters were solved using a script to resample the textures exported from Fuse to achieve realistic quality.

![Child avatars designed in Adobe Fuse, rigged in Mixamo's web service, and imported into a Unity scene for users to embody.](image)

*Figure 13. Child avatars designed in Adobe Fuse, rigged in Mixamo's web service, and imported into a Unity scene for users to embody.*
5.3.3 Body Tracking

The platform requires a high-resolution awareness of the user's body in order to establish motor-visual synchrony and create the body ownership illusion. Many technologies exist that could allow this, such as Microsoft's Kinect or body suit alternatives. However, most available solutions are cumbersome. Body suits in the market require the user to buy a specific size to adjust to their bodies, and having the user put on a whole body suit every time they will be immersed in the virtual environment becomes a tedious chore. Our goal with Inner Child was to reach a solution that would give us good resolution for tracking while minimizing the complexity for the user. Vision based alternatives, such as the Kinect, were discarded due to the inability (without a more complex rig) to provide full 360 degrees of freedom to user's. We wanted users to be able to move freely within the virtual space.

Inner Child makes use of HTC Vive Trackers, which function similarly to the Vive's control, providing high resolution tracking for a single point in space. The headset tracks the user's head, while the two controllers provide the position of the user's hands. Two trackers were employed to track the participant's feet (one on each foot), using a strap that wraps around the sole of the shoe. A belt is used to wrap around the user's waist, using Velcro straps to adjust for size. Additionally, two more trackers could be used to track the direction of the elbows. These trackers were optional to the setup, and were not required to achieve the body ownership illusion, but would enhance the avatar's synchrony with the user's movements.

In order to minimize the number of trackers required to operate the system, we used inverse kinematics provided by Final IK, to actuate the avatar's limbs more realistically. Inverse kinematics is the process of obtaining missing information based on pre-defined constraints of the joint parameters of end-effectors. In other words, for the purpose of Inner Child, we want to approximate how the arms or legs bend based on the position of the end-effector (i.e. hands and feet). Instead of using too many trackers that hinder the usability of the system, we narrow the tracking points to those that maximize the amount of information needed to fill in the blanks. This allowed us to achieve very realistic depiction of full body movements without trackers on the knees, elbows, chest, or shoulders.
However, without trackers on the knees or elbows, the target angle at which the joint bends is not calculable.

The direction towards which these joints bend was approximated by pairing a game object in Unity to the tracker's position. This made the direction, which we denominate as the bend goal, paired to the rotation of the tracker. In the case of the foot trackers, we project the bend goal upwards from the foot tracker by half the distance between the waist and the foot tracker. As a result, when the knee bends, it does so in the direction where the instep (i.e. upper part of the foot) aims at. Because the foot is limited in the degrees of freedom by which it bends to the sides at the ankle, this approximation gives a very realistic feel that the knee direction is accurately represented by the virtual body. A similar approach is employed for the approximation of the elbows. However, because the hand has more freedom for rotation at the wrist than the foot at the angle, certain movements can make the resulting bend goal seem artificial and break the illusion. For this reason, we employed two additional trackers for the elbows that are optional in the experience.

![Figure 14. Suits like the Perception Neuron allow fine grained body tracking in virtual reality at the cost of cumbersome usability. Our system only employs three trackers to achieve body tracking with a minimal number of points on the user's body.](image-url)
During the calibration phase the user is asked to stand in a T-pose to measure height. Height is measured using the foot trackers position and HMD position, and subsequently used to scale the avatar’s body and environment.

5.3.4 Three-Step System

Inner Child’s system divides the embodiment illusion into three steps: calibration, conditioning, and persistence. Calibration involves the way we adjust the avatar and the environment to the user’s body. Conditioning entails the use of a methodical approach of establishing visual-motor synchrony and rapport with the new body. Finally, persistence comprises cues and re-affirmation of the change after it has been initially established.

5.3.4.1 Calibration

The calibration stage consists of two processes. The first step in the calibration stage is identifying all the trackers active in the scene and assigning them to a particular limb. Vive trackers are assigned dynamically in the scene. This means that a tracker can have a variable reference in the scene, which in turn requires a dynamic allocation method to each limb. Because the participant is standing in a T-position when the calibration begins, certain assumptions can be made about the relative position of each tracker with respect to the user’s body. A plane of reference is created using the headset’s forward vector as a
plane normal. We then project the position of all the Vive trackers onto this plane. The relative position of objects in this plane is evaluated to determine the tracker position with respect to the user's body. The lowest two trackers are assigned to the feet, and the upper two trackers assigned to the hands (four if elbows trackers are included). The remaining tracker is automatically assigned to the waist. A similar process is used to determine which tracker to assign as left or right hand/foot.

The second process in the calibration stage consists of measuring the user's height, which we calculate as the distance between the feet trackers and the HMD. The same distance is measured on the target avatar. A scaling factor is obtained by dividing the user's height (which is presumably larger due to the avatar being a child) by the avatars height. The avatar is then scaled by this factor to make it as tall as the user. The world is subsequently scaled up to keep the relationship between avatar and world intact.

5.3.4.2 Conditioning

During the conditioning stage of the Inner Child experience, participants are tasked with completing a systematic series of poses that involves their whole body to establish rapport with the new body they have taken ownership over. As opposed to other virtual ownership experiences in the literature, which often ask participants to move about and explore the control they have over their new body in an unstructured format, our rationale for subsequent experiments was that we required a way to create this rapport in a way that was uniform for all participants. This activity, which is meant to be performed in front of a virtual mirror, involves the user's hands and feet in activating a series of orbs. When touched with the avatar's hand the orb lights up green, indicating that it is active. When all the visible orbs in the scene are activated simultaneously, the pose is completed and the next pose is queued.
The Inner Child activity makes use of twenty poses that increment in difficulty, beginning with the use of a single limb and extending to the use of all four limbs at the same time. The poses that a participant must perform are strategically selected so that users have to (1) look at their own body for reference, (2) look at the mirror to synchronize their movements, and (3) understand the dimensional constraints of their body. Point 1 is achieved by placing orbs at distinct distances in the z-axis. Since the mirror only gives the participant a cue of the position of an orb in two dimensions (up-down/right-left), participants must look down or at their body to ensure they adjust the position correctly in the third dimension (forward-backwards). Point 2 is a result of placing multiple orbs at the same time. Once the user places a limb on one of the orbs, he must maintain it there while he/she actuates the remaining orbs. This requires the participant to maintain a visual reference of the whole body during the execution of multiple orb poses. Finally, point 3 results from the nature of the selected poses, which often ask the user to keep their feet in a position while actuating an orb that is at the limit of their reach with their new body. This is especially important in the case of a child’s body, in which a participant is often limited by the reach of their arms or height.

5.3.4.3 Persistence

The final step involves introducing elements in the scene that passively reinforce the already established illusion. In contrast with calibration and conditioning, the persistence
stage is not discreet as much as it is recurring. Elements in the virtual classroom reinforce the feeling of embodiment throughout the learning experiences. This primarily involves the use of a mirror or reflective surface at strategic locations in the scene, as well as cues for comparison of scales. In the Inner Child experience this is achieved by positioning two adults in the periphery of the participant’s view, as well as a peer child student in front of them. This visually reinforces the sense of scale between the user’s body, the child body of a peer, and the significantly bigger adult bodies nearby. Additionally, several objects related to the scene are positioned behind the student peer and adult avatars, establishing a clear distinction from the user’s point of view of the scale between the environment and their own body. Finally, during the learning portion of the Inner Child experience, mirrors in the periphery show the user’s body, anchoring the user’s relationship with their virtual body even if they are not employing it in the learning activity.

5.3.5 Voice
The Inner Child experience involves a feedback loop that alters the participant’s voice. A child-like analogue is produced by increasing the pitch and shifting formants of the user’s natural voice, which is captured using a microphone connected to the Vive setup, altered using Garage Band, and fed back through noise cancelling headphones. In this work, our focus was not on creating a natural child voice changer. This is a challenging problem which requires understanding of each user’s vocal features, which is beyond the scope of our research. Our goal was to achieve an artificial voice which didn’t break the body ownership illusion throughout the experience. We achieved this by (1) equalizing (2) slightly increasing pitch, and (3) shifting the formant of the user’s voice using filters. Additionally, we use a noise filter to cut-off any background noise that may be captured by the microphone.

In this chapter, we have presented the development of Inner Child, an implicit embodied language learning platform that aims to decrease the subjective age of users to enhance the way they approach language learning. In the following chapter, we formally evaluate Inner Child and its ability to impact language education.
Chapter 6

IMPLICIT EMBODIED LEARNING STUDY

We performed a similar study to the one presented in Chapter 4, to study the effect of implicit embodied learning in VR. An experiment was conducted with 60 participants to understand the effects of illusory ownership of a child’s body in virtual reality while approaching a simple vocabulary learning task. The experiment consisted of three groups: learners with no virtual body (No Body), learners with an adult virtual body (Adult), and learners with a child virtual body (Child). Each group is composed of 20 participants randomly assigned to an experimental condition. All participants were recruited from campus and awarded a $10 Amazon gift card upon completion of the procedure.

6.1 Vocabulary Acquisition Study Design

Participants were recruited through email. Interested students enrolled in 1-hour slots of time to visit the lab and engage in the virtual reality learning experience. Upon arrival, participants were briefed on the experimental procedure and the required consent was obtained. Afterwards, subjects in all groups were submitted to a pre-test, which contained 25 words in Spanish and required the subject to fill out the corresponding word in English. This pre-test allowed us to gauge the participant’s initial knowledge (if any) of the words that they will learn throughout the experience.

After taking the pre-test, the participant was fitted with the HTC Vive trackers: one on each foot, one on each elbow, and one around the waist. When the trackers were correctly placed, participants entered virtual reality. Initially, subjects in all groups were disembodied in front of a large mirror. Subjects were able to see the trackers attached to their bodies, along with the two Vive controller models. Instructions displayed on a floating panel next to the mirror asked for participants to pull the control’s trigger (in order to teach them how to advance instructions and interact with the virtual environment). Afterwards, participants were asked to stand in a T-pose (feet together and arms stretched to each side)
and pull the trigger on the Vive controller. This calibrated the system and the environment to make it independent of the participant’s height.

Subjects in the adult embodied group appeared in the virtual classroom, with the body of an adult that matches their selected gender. Subjects in the child embodied group appeared in the virtual classroom, with the body of a child that matched their selected gender. At this point, the control and tracker models were deactivated. The disembodied group appeared in the virtual classroom, with no body. The Vive trackers and controllers were still visible to give them a sense of presence, and allow them to identify their location in front of the mirror. After embodying their respective avatars (or lack thereof), subjects were instructed to look at their virtual self in a mirror.

The conditioning phase consists of a set of 20 poses that the participant had to match. This allowed us to cement the embodiment illusion through motor visual synchrony in a controlled way, where every participant engaged in the same prescribed movements. The subjects performed a pose by matching “target orbs” that appear in the scene with their hands and feet. Whenever a target orb was touched, it changed color from white to green, to indicate it had been activated. Similarly, if touch was disengaged the target orb would revert to its inactive state, signaled by the white color. A pose was considered complete once all the target orbs were activated simultaneously.

The learning phase consisted of the participants observing a set of word pairs that cycle on top of the mirror placed in the virtual classroom. The mirror is positioned so that the participant can focus on the vocabulary words while retaining a peripheral view of their virtual avatar. Each word pair consists of the target word in Spanish and its translation in English. The training consisted of 20 word pairs that will cycle automatically, presenting each word pair twice for fifteen (15) seconds, totaling forty-five (45) seconds of exposure for each new vocabulary word. Although a voice component was implemented for the Inner Child experience, participants were instructed not to talk during the learning phase, as we wanted to isolate the effect of the virtual body’s appearance, and an altered voice would create an unnecessary confound.
Right after instruction, all groups were prompted to complete a post-test outside of VR. The post-test contained the same words as the pre-test. The post-test also included qualitative questions regarding the virtual reality experience (Appendix A). The post-test was administered through the same medium as the pre-test. All groups were prompted to complete two delayed post-tests (one week and two weeks after the post-test was administered). Once more, the delayed post-tests contained the same words as the immediate post-test. Figure 22 depicts a visual workflow of the described procedure for all experimental conditions. Out of the 60 participants, 58 completed the one week delayed test, and 57 also completed the two weeks delayed test.

<table>
<thead>
<tr>
<th>Group</th>
<th>Age</th>
<th>Gender</th>
<th>Lang. Spoken</th>
<th>Lang. Studied</th>
<th>VRExposure (7-point Likert)</th>
<th>Lang. Difficulty (7-point Likert)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Body</td>
<td>28.05</td>
<td>5.59</td>
<td>50%</td>
<td>50%</td>
<td>2.55 1.05</td>
<td>2.55 1.54</td>
</tr>
<tr>
<td>Child Body</td>
<td>29.25</td>
<td>6.52</td>
<td>55%</td>
<td>45%</td>
<td>1.83 0.94</td>
<td>2.70 2.08</td>
</tr>
<tr>
<td>Adult Body</td>
<td>28.16</td>
<td>4.10</td>
<td>31.6%</td>
<td>69.4%</td>
<td>2.61 1.55</td>
<td>2.47 1.84</td>
</tr>
</tbody>
</table>

Table 1. Participant age, gender distribution, number of languages spoken (Lang. Spoken), number of languages studied in the past (Lang. Studied), exposure to VR systems (VRExposure), and self-reported difficulty when learning a new language (Lang. Difficulty), divided by body groups.
6.2 Results

One-Way ANOVA confirms that participants in the three body groups showed no difference in terms of age (p=0.75), self-reported difficulty at learning a language (p=0.13), self-reported exposure to virtual reality systems (p=0.93), number of languages spoken (p=0.083), and number of languages studied (p=0.41). The individual means and standard deviations of the participant’s responses to these measures can be found in Table 1. No differences were found between all body groups in terms of prior knowledge, with only four participants who were able to identify a single word in the pre-test.

One-Way ANOVA shows no significant differences in performance based on body condition during tests administered immediately (p=0.13), one week (p=0.24), and two weeks (p=0.39) after the exposure to the virtual reality experience. The Child group performs immediately better than the no body condition and adult groups (Figure 23) but these differences were not found to be significant. A similar trend is observed one week after, with participants in the Child and Adult group outperforming those in the No Body group, but no significant difference is observed. Despite no significant differences observed by body condition, we are interested in understanding the effects of subjective age change and embodiment, which can vary within-groups (e.g. not everybody who embodied a child, felt like a child or felt ownership of the child body).

![Test Results by Body Group](image_url)

*Figure 23. Test results by body group immediately, one week, and two weeks after experience.*
Figure 24. Boxplot of embodiment measures by body group obtained from exit surveys.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>MyBody</td>
<td>How much did you feel that the virtual body you saw when you looked down at yourself was your own body? (Despite it might not look like you)</td>
</tr>
<tr>
<td>Mirror</td>
<td>How much did you feel that the virtual body you saw when you looked at yourself in the mirror was your own body? (Despite it might not look like you)</td>
</tr>
<tr>
<td>TwoBodies</td>
<td>How much did you feel as if you had two bodies?</td>
</tr>
<tr>
<td>LikeMe</td>
<td>How much did you feel that your virtual body resembled your own (real) body in terms of shape, skin tone, or other visual features?</td>
</tr>
</tbody>
</table>

Table 2. Variables used to characterize illusory ownership using a 7-point Likert Scale in exit survey.

Four variables were used to characterize the success of the embodiment illusion across body groups (see Table 2 and Figure 24). Kruskal-Wallis analysis found no significant differences in MyBody ($p=0.60$), Mirror($p=0.37$), or LikeMe ($p=0.391$) variables across body groups. However, the test indicates a difference exists amongst groups in the TwoBodies control variable ($p=0.02$). Post-Hoc analysis reveals significant differences between the Child group and both No Body and Adult group responses for TwoBodies ($p<0.05$).

We use variables MyBody and Mirror to create a composite measure of embodiment by averaging the participant's responses. The resulting variable (Embodiment) is used to characterize participants who felt the illusion of ownership (High Embodiment), and those who did not (Low Embodiment). The method for dichotomization is a median split which
resulted in participants with a measure higher than 4 (neutral on the 7-point Likert Scale) characterized as feeling embodied in their avatars, while participants equal or lesser to 4 being characterized as not feeling embodied. Although dichotomizing a continuous variable is generally not recommended, it is acceptable under some circumstances, such as when there exist two different groups. We recognize that there are two groups for our analysis, those who were under the illusion that their virtual avatar was their body and those who did not. As such, a dichotomization of the variable can yield insight for the purpose of this analysis. In this case, the median split coincides with the logical interpretation of the 7-point Likert scale data (i.e. greater than 4 are in agreement that the body belongs to them, while lower or equal than 4 do not feel a sense of embodiment). Two-Sample t-tests show there is no significant differences between both groups in terms of age (p=0.88), self-reported difficulty at learning a language (p=0.074), exposure to virtual reality systems (p=0.60), number of languages spoken (p=0.84), and number of languages studied (p=0.17).

<table>
<thead>
<tr>
<th>Group</th>
<th>Age</th>
<th>Gender</th>
<th>Lang. Spoken</th>
<th>Lang. Studied</th>
<th>VRExposure (7-point Likert)</th>
<th>Lang. Difficulty (7-point Likert)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Embodiment</td>
<td>28.36</td>
<td>5.15</td>
<td>28%</td>
<td>72%</td>
<td>3.12</td>
<td>1.64</td>
</tr>
<tr>
<td>High Embodiment</td>
<td>28.59</td>
<td>5.73</td>
<td>59%</td>
<td>41%</td>
<td>2.29</td>
<td>1.24</td>
</tr>
</tbody>
</table>

Table 3. Participant age, gender distribution, number of languages spoken (Lang. Spoken), number of languages studied in the past (Lang. Studied), exposure to VR systems (VRExposure), and self-reported difficulty when learning a new language (Lang. Difficulty), divided by embodiment group.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repeat</td>
<td>How inclined would you be to learn new vocabulary through this new virtual reality method?</td>
</tr>
<tr>
<td>Engagement</td>
<td>How engaging did you find the technology enhanced learning experience?</td>
</tr>
<tr>
<td>Intuitiveness</td>
<td>How intuitive did you find the learning experience to be?</td>
</tr>
</tbody>
</table>

Table 4. Variables used to characterize subjective metrics using a 7-point Likert Scale in exit survey.
Two-Sample t-test shows that participants in the High Embodiment group (M=13.7, SE=0.69) performed significantly better, remembering more words on average than those in the Low Embodiment condition (M=10.64, SE=1.04), immediately after being exposed to the virtual reality experience (p=0.013). No significant differences were found in performance a week (p=0.28) or two weeks (p=0.30) after (Figure 26). The participant's attitude towards the experience was significantly better for participants in the High Embodiment group (Figure 25). Variable *Repeat*, which acts as an indicator of willingness to participate in a similar learning experience in the future is significantly higher for the High Embodiment group (M=3.79, SE=0.31) than for the Low Embodiment group (M=2.56, SE=0.32) (p=0.008). A similar effect is observed in engagement, where participants that experience high body ownership (M=3.7, SE=0.26) report significantly higher measures of *Engagement* than those who felt low body ownership (M=2.8, SE=0.34) (p=0.037). There was no significant difference in self-reported intuitiveness regarding the experience between groups (p=0.29).

![Subjective Metric Scores by Embodiment Condition](image)

*Figure 25. Subjective metric scores divided by embodiment condition show a significant increase in Engagement and Repeat for subjects who felt a high degree of ownership.*
Figure 26. Test results by embodiment group immediately, one week, and two weeks after experience show a significant, immediate increase in performance for subjects with higher levels of embodiment.

Figure 27. Within-Group analysis of embodiment by body condition shows significant differences only occur for subjects with no humanoid body.
Figure 28. Significant difference in the number of words recalled are also observed in within-group analysis of subjects with no humanoid body, one week and two weeks after exposure.

Splitting participants with high embodiment and low embodiment by body group shows no immediate differences between subjects in the Child (p=0.834) and Adult (p=0.28) condition caused due to embodiment. This trend holds for one week (Child, p=0.26; Adult, p=0.891) and two weeks (Child, p=0.16; Adult, p=1.0) after the learning experience. However, subjects within the No Body group in the Low Embodiment condition perform significantly worse immediately (p=0.008), one week (p=0.026), and two weeks (p=0.019) after exposure than those in the High Embodiment condition (Figure 27 & 28).

Engagement and Repeat show an increase in the Adult and Child conditions based on body ownership (Figure 29). Participants in the Adult condition who felt a high degree of body ownership felt significantly more engaged (M=4.5, SE=0.48) in the learning activity than those who felt a lower degree of body ownership (M=2.33, SE=0.47) (p=0.005). Similarly, subjects in the Adult condition who felt ownership (M=4.6, SE=0.60) show a significantly higher Repeat measure than those who did not (M=2.56, SE=0.60) (p=0.02). Similar trends are observed in the Child condition, where participants feeling ownership show higher measures of Repeat (p=0.05) and Engagement (p=0.21) that those with low ownership. Subjects in the Nobody condition, show no difference in Engagement (p=0.698) or Repeat measures due to illusory body ownership (p=1.00).
In order to understand the effects of making subjects feel younger on the learning outcome, we collected measures regarding the perceived age of the virtual avatar (PerceivedAge), how old the participants felt during the experience (FeltAge), and how old the participants were (MyAge) (Figure 30). We define an illusory age difference as the percentage change in age, to normalize by age throughout our analysis.

\[
FeltChange = \frac{MyAge - FeltAge}{MyAge}
\]
Avatar's Perceived Age Versus Participant's Felt Age

The perceived age of the avatar was positively correlated with the subjective age a participant felt throughout the Inner Child experience.

The mean Perceived Age of the avatar in participants in the Adult (M=19.28, SE=0.89), Child (M=10.30, SE=0.89), and No Body (M=18.23, SE=1.71) conditions show that the adult and child avatars were successfully perceived as their intended age group. One-Way ANOVA detects a significant difference in Perceived Age (F(2,56)=15.699, p<0.001), Felt Age (F(2,56)=6.675, p=0.003), and Felt Change (F(2,56)=7.034, p=0.002) amongst body groups. Post-hoc Tukey-Kramer tests showed that the perceived age of the virtual avatar was significantly lower between the Child and Adult group (p<0.05); post-hoc tests also show this significant change holds between the No Body and Child groups (p<0.05). In the case of Felt Age and Felt Change, these differences were only significant between the Child and Adult groups (p<0.05). The perceived age of the virtual avatar was significantly correlated with the change in subjective age reported by the participant (r(57)=0.648, p<0.001). I.e., younger perceived age was correlated with feeling a younger age (Figure 31).

Analysis of Felt Age and Felt Change across groups reveals that participants in the Adult and No Body group still felt a change in subjective age. Likewise, some subjects in the Child
group felt no change in their subjective age. We divide participants into two groups, those who felt a change in subjective age (Younger), and those who didn’t (No Change). This dichotomization of the *FeltChange* variable is justified by the clear definition of two distinct groups, namely those who felt change (34 participants) in their age and those who didn’t (25 participants). Participants with 0% *FeltChange* were assigned to the No Change group and participants with positive *FeltChange* were assigned to the Younger group. Variables *FeltYoung*, *FeltLike*, and *FeltChild* support this dichotomization with participants in the Younger group feeling significantly younger (p<0.001), identifying in larger proportion as children (p=0.002), and affiliating more strongly with the feeling of being a child (p<0.001) than subjects in the No Change group.

![Boxplot showing subjective age measures for subjects divided into the Younger and No Change groups.](image)

**Figure 32.** Boxplot shows the subjective age measures for subjects divided into the Younger and No Change groups.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>FeltYoung</em></td>
<td>How much younger did you feel?</td>
</tr>
<tr>
<td><em>FeltLike</em></td>
<td>During the experiment, you felt like: (a) Child, (b) Adult</td>
</tr>
<tr>
<td><em>FeltChild</em></td>
<td>How much did you feel like a child?</td>
</tr>
</tbody>
</table>

**Table 5.** Qualitative measures used in the study to assess subjective age change in participants.
Figure 33. The number of words recalled is higher for participants in the Younger condition immediately, one week, and two weeks after exposure. These changes are more prominent for participants with humanoid virtual bodies.

Two-Sample t-tests show no difference in exposure to virtual reality (p=0.52), number of languages spoken (p=0.61), number of languages studied in the past (p=0.81), or difficulty learning a new language (p=0.308) between the Younger and No Change groups. There is, however, a significant difference in the mean age of participants between groups; subjects in the Younger condition (M=27.2, SD=5.58) were slightly younger than those in the No Change condition (M=30.2, SD=4.86). No difference in embodiment was found between participants who felt younger and those who didn’t (p=0.40).

Subjects who felt younger (M=13.32, SE=0.76) recalled more words on average than those who felt no change (M=11.16, SE=0.99) in their subjective age immediately after the experience (p=0.084). This difference is also observable one week after participating in the experience, with participants in the Younger group (M=7.61, SE=0.88) outperforming those in the No Change group (M=5.46, SE=0.83) (p=0.09).

In order to understand the effects of subjective age change on retention, we define the metric percentage of words lost, which is the ratio of the amount of words forgotten between the immediate test and the delayed test:

\[
\% \text{ words lost} = \frac{\text{immediate score} - \text{delayed score}}{\text{immediate score}} \times 100
\]
Figure 34. Percentage of words lost is significantly lower for participants who identified with feeling like a child through the experience for both embodied and non-embodied subjects.

Subjects in the Younger group (M=44.58, SE=4.75) displayed a lower mean percentage of words lost that those in the No Change group (M=53.50, SE=4.85) (p=0.20). Moreover, participants who reported feeling like a child (M=40.21, SE=6.05) display a significantly lower mean percentage of words lost than those who report feeling like adults (M=53.83, SE=3.87), as measured by the FeltLike variable (p=0.05) (Figure 34). Engagement and Repeat are also affected by FeltLike, where participants feeling like children feeling more open to repeating the experience (p=0.081) and significantly more engaged (p=0.041).

These differences become more prominent when observing participants who had humanoid bodies (both Adult and Child conditions). The mean number of words recalled by participants in the Younger embodied group (M=14.25, SD=0.85) is higher than those of the No Change embodied group (M=11.67, SD=1.09) (p=0.067). A week after, the performance of the Younger embodied group (M=8.70, SE=1.08) is significantly higher than the No Change embodied group (M=5.50, SE=0.91) (p=0.047). The mean percentage of words lost is significantly lower for embodied participants that report feeling like a child (M=36.50, SE=7.65) than those who felt like adults (M=53.49, SE=4.00) (p=0.04), which is also the case for the Younger embodied (M=40.5, SE=5.75) and No Change (M=56.6, SE=4.24) embodied groups (p=0.05).
Figure 35. Changes in subjective age were significantly correlated with embodying the avatar of a child, negatively correlated with embodying the avatar of an adult, and not correlated for subjects with no humanoid body.

Taking into account Younger/No Change participants and Younger/No Change embodied subjects showed a significant higher proportion of subjects in the child group in both the Younger and Younger embodied groups (Figure 35). This shows a positive correlation between being embodied in a child avatar and feeling younger ($r(57)=0.414, p=0.001$), a negative correlation between feeling younger and embodying an adult avatar ($r(57)=-0.353, p=0.006$), and no correlation between feeling younger and having no body in virtual reality ($p=0.62$).

6.3 Analysis

We divide the analysis of the aforementioned results into two claims: having a virtual body improves learning experiences supported by virtual reality and making that avatar a child can impact the mindset of participants such that it improves the overall learning outcome.

6.3.1 Impact of Having a Virtual Body

Our first claim posits that having a virtual body is beneficial for learning activities as a countermeasure of the feeling of disembodiment. The measure of embodiment utilized in this analysis was the average of how much the participants felt ownership of their bodies when looking at a virtual mirror and downwards at themselves. The dichotomization of this measure into subjects who felt ownership (high embodiment) and those who didn’t feel
ownership (low embodiment) showed significant differences in learning outcomes across and within groups. However, there was a feeling of embodiment in equal measures across all body groups, which suggests that showing the controls and trackers floating in space is sufficient to induce the feeling of embodiment. Since results indicate the LikeMe variable has a higher variance for participants in the Nobody condition, it can lead to lesser measures of embodiment, albeit not significant for the purpose of this study.

Despite there is no evidence in the data to support that having a humanoid body increases embodiment in comparison to a non-humanoid representation (floating trackers), there is a difference in the effect of the learning outcome when a humanoid body is not present. Within group analysis of performance immediately and a week after exposure shows no significant differences between participants who felt embodied or disembodied in the Adult and Child groups. However, in the No Body condition, participants who felt disembodied significantly underperformed those who felt a degree of ownership. In other words, although the presence of a virtual humanoid body does not affect the measure of embodiment, it is a deterrent of the negative effects of feeling disembodied.

The effects of ownership on the immediate and delayed recall tests can be attributed to two different reasons. Having a body that resembles a human being can result in a heightened sense of presence, which has been associated with improved recall performance in the literature [94, 95, 96]. This would explain why the percentage of words forgotten between immediate and delayed tests is not affected by the measure of embodiment. In a sense, having a virtual body generates more attentional focus and allows participants to learn more words immediately after exposure by enhancing memory. However, the retention or memorability is not affected by the body itself. An alternative hypothesis is that, because the words in the learning experience are verbs, a virtual representation of the body stimulates motor regions in the brain, which in turn creates a more memorable encoding associating body parts with the word itself.

Participants in the Child and Adult conditions who felt low ownership of their virtual bodies did not suffer from reduced recall scores. In a sense, the body grounds the participant’s attention, despite them not feeling embodied. A user study showed that “the
sole effect of congruent visuoproprioceptive cues, provided by a high degree of spatial overlap between the physical and virtual bodies, is a sufficient condition for inducing a full body ownership” [18]. Having a virtual body while the participant is learning the words in the virtual mirror creates a higher degree of overlap between physical and virtual bodies, thereby grounding presence despite lower measures of ownership. This supports the hypothesis that having a virtual human body can preemptively ensure that lesser sense of ownership does not significantly impact language learning experiences in virtual reality.

Self-reported engagement and willingness to participate in language learning activities of the same kind was significantly higher for subjects who felt embodied. We acknowledge that the scores for engagement and repeat were both low on the Likert scale measures, however, this is not unexpected given the learning portion of the activity did not involve any memorable experience (as no participant figured out the purpose of the experiment which was to understand the effect of the virtual body). Most participants complained that learning a word by simply observing it in virtual reality was not very entertaining. Despite this, we are interested in the difference between participants with a higher degree of ownership and those with lower degrees of ownership. Within group analysis showed that higher degree of ownership significantly increased engagement and repeat only in groups that had a humanoid virtual body, whereas there was no difference in these measures for participants in the No Body condition. This supports the notion that having a virtual body during learning activities can not only affect the learning outcome, but can positively impact a student’s attitude towards the modality as well.

6.3.2 Impact of Embodying a Child Avatar
Participants who experienced a change in subjective age outperformed participants who didn’t both immediately and a week after exposure to the virtual reality experience. In the Inner Child experience, we were successful inducing a subjective age change for participants in the Child and Adult groups. This can be observed by the significant differences in FeltAge between the Child and Adult group. The variance observed in FeltAge on the disembodied participants suggests that in absence of a physical body, subjects delegated their subjective
age to the environment, which aligns with findings from Langer [64] and Wilson [21], which posit cognition is leveraged on the world. Correlations found between perceived age of the virtual avatar and the subjective age felt by participants suggest that the effect of change in subjective age is induced through the avatar's visual features. The composition of the Younger group showed a positive correlation for a child body, a negative correlation for an adult body, and no correlation for participants with no body. These findings support the notion that virtual reality can be leveraged to induce changes in subjective age by illusory ownership of a younger body.

The effects of induced subjective age become more apparent when controlling for participants with virtual bodies. The difference in performance both immediately and delayed recall tests is significant for those who felt younger and had a virtual body than for those who did not feel a change in subjective age. This difference is stronger a week after the tests were administered. Moreover, there was a significant decrease in percentage of words lost for participants who felt younger and had a body. This suggests that, unlike the effects of embodiment which only resulted in immediate increase of words recalled, feeling a change in subjective age has a positive effect in the retention of new material.

A younger subjective age has been found to enhance recall and delayed recall in older populations [97]. Moreover, induced changes in subjective age have shown to induce effects in both cognitive [98] and physical tasks [99]. Feeling younger has been shown to enhance self-efficacy in adults, which can in turn significantly affect the outcome of a task [100, 101]. That is, our belief in our abilities impacts our performance in them. Participants embodying adult avatars might have felt more pressure to succeed at the task, while those feeling like children approached the task with an open or more relaxed mindset, consequently outperforming subjects who felt like adults during the experience.
CONCLUSION

In this thesis, we have outlined two different modalities for embodied learning. Explicit embodied learning directly uses the body to encode or represent information as a tool to enhance learning outcomes. Implicit embodied learning uses the body’s appearance to alter the student’s self-perception and consequently change the way students learn. Virtual reality is presented as a platform that affords both modalities. Explicit embodied learning is supported by the platform’s capability to spatially understand the user's body. Implicit embodied learning is supported by the platform’s capability to induce strong illusions of alternate body ownership. Two systems were implemented to test the effectiveness of explicit (Words in Motion) and implicit (Inner Child) embodied language learning in VR.

A study, which involved 60 subjects recruited from the university campus, tested the effects of the Words in Motion kinesthetic approach to learn 20 new vocabulary words. Participants were divided into three conditions: a text-only condition outside of virtual reality, learning kinesthetically in virtual reality, and observing words inside virtual reality without performing actions. Results showed that participants in the text-only condition initially outperform virtual kinesthetic and virtual non-kinesthetic subjects for equal exposure time to the material. Both groups in virtual reality performed similarly immediately after exposure. However, a week after exposure, subjects in the virtual kinesthetic group significantly outperformed those in the virtual non-kinesthetic group and showed no difference to participants in the text-only group. Moreover, the number of times a word was remembered was directly correlated to the number of times the action associated with that word was performed both in immediate and delayed evaluations. In other words, performing actions in VR has a positive effect on the retention of new vocabulary.

Although participants in kinesthetic and text-only groups showed no significant difference a week after they were exposed to the material, the retention rate was significantly higher for subjects in the virtual kinesthetic condition. The fact that the retention rate between text-only and virtual non-kinesthetic are not different further supports the
hypothesis that this effect is due to the kinesthetic component, and support the virtual kinesthetic modality as more effective for retaining new vocabulary words in comparison to non-kinesthetic VR and text-only modalities.

The findings in this experiment support the hypothesis that virtual reality can benefit from explicit kinesthetic elements to enhance language learning activities. However, they also highlight that virtual reality kinesthetic learning is characteristically different from real-world and non-immersive technology enhanced kinesthetic learning, where exposure has been shown to result in higher immediate learning gains. Nevertheless, virtual kinesthetic learning shows the same enhanced retention effect as real-world and non-immersive analogues. Given the positive effect on retention, this work suggests that with additional exposure and conditioning to the effect of “novelty” of VR, kinesthetic language learning in virtual reality can positively impact language education.

The effects of implicit embodied learning were explored with 60 participants using the Inner Child platform. Our hypothesis was two-fold: having a virtual body impacts language learning in virtual reality, and employing a child avatar can enhance the student’s performance by inducing a change in subjective age through illusory ownership of the younger body. Participants were split into three groups that corresponded to the respective avatar they took ownership of during the learning activity: No Body, Child, and Adult. Their task was to learn 20 vocabulary words by observing them in the virtual environment while standing in front of a mirror. Prior to exposure to the vocabulary words, participants performed poses to create rapport with the body they were inhabiting in the virtual simulation. Their learning was evaluated through tests that were administered before, right after, a week, and two weeks after exposure to the virtual group condition.

Results showed that participants who felt embodied performed significantly better immediately after exposure. Within group analysis revealed that there was no difference between subjects in groups that had a humanoid virtual body (i.e. Child and Adult groups), but a significant difference with the No Body condition. A week after, across all 3 groups, no difference is observed. However, within group analysis also revealed that participants in the No Body condition who felt higher degree of ownership still performed significantly better
than those who felt low degree of ownership. These effects can be attributed to a higher degree of presence, which is associated with enhanced immediate and delayed recall. Results support the claim that a humanoid body is not necessary to create illusory ownership and its benefits on recall. However, having a virtual body can deter the effects of feeling low body ownership by grounding presence through increased overlap between virtual and physical body. Participants with humanoid virtual bodies who felt a high degree of ownership also reported significantly higher levels of engagement and willingness to repeat the learning experience. This supports the hypothesis that having a humanoid virtual body can impact learning outcomes in virtual reality, as well as impact a student’s attitude towards learning experiences.

The Inner Child experience was successful in inducing a subjective age change in participants, with subjects in the Child group reporting a significantly lower subjective age than those in the Adult and No Body conditions. This change was found to be positively correlated with the perceived age of the virtual avatar that the subject embodied. Participants who felt a change in their subjective age (i.e. felt younger), outperformed those who felt no change in subjective age immediately, one week, and two weeks after. Moreover, participants who reported feeling like children had a significantly lower percentage of words lost between immediate and delayed tests. These effects become more prominent when comparing subjects who had humanoid virtual bodies.

Amongst subjects who felt younger and had a humanoid virtual body, the change in subjective age also significantly enhanced the retention of vocabulary words between immediate and delayed tests. Having the virtual body of a child was positively correlated with a change in subjective age, whereas having the virtual body of an Adult was negatively correlated with changes in subjective age. And having no humanoid body had no positive or negative correlation. This supports our hypothesis that embodying a child avatar in virtual reality can positively impact language education by inducing a subjective age change, which is associated with changes in mindset that can improve a student’s performance in learning activities.
This work explored implicit and explicit embodied learning as isolated modalities. However, implicit and explicit approaches could be combined to create powerful learning experiences. Words in Motion could be combined with Inner Child, allowing students to learn a second language using action (explicit), under the illusion of being children (implicit). We hypothesize that some of the benefits on learning of these two approaches are orthogonal in nature. Words in Motion exhibited an enhanced retention, but subjects using the system underperformed immediately after exposure. Conversely, the effects of having a virtual body and presence in the Inner Child experience show an enhanced immediate acquisition of words, without affecting retention. This is an example of how explicit and implicit modalities can be combined to enable powerful learning in VR.

In synthesis, two prototypes were built to explore explicit and implicit modalities of embodied learning, focusing on language learning because theories of embodied cognition postulate a relationship between language and the body. Words in Motion enables explicit embodied language learning by recognizing student actions with objects in virtual spaces. Executing an action with a virtual object results in the target second language word floating momentarily over the object, which enables the encoding of the target word in association with the performed action. Inner Child enabled implicit embodied language learning by creating the illusory ownership of a younger body, effectively decreasing the subjective age of students to enhance their mindset during learning tasks. Results from both explicit and implicit embodied applications show that virtual reality as a platform holds potential for language education that makes use of the body to enhance language instruction.
APPENDIX

A. Passive Recall Test

The following test was used to assess the participants in both explicit and implicit embodied language learning studies presented in Chapter 4 and Chapter 6.

Each item in this form will prompt you with a word in the target language (Spanish). There are a total of twenty-five (25) items. For each item, please provide the English translation of the word in the field below. If you do not know the translation of the word, leave the field blank.

* Required

Email address *

Your email

cornillar

Your answer

Embruajar

Your answer

Espetar

Your answer

Rapar

Your answer

Lacerar

Your answer

Never submit passwords through Google Forms.

NEXT
Instructions

Each item in this form will prompt you with a word in the target language (Spanish). There are a total of twenty five (25) items. For each item, please provide the English translation of the word in the field below. If you do not know the translation of the word, leave the field blank.

pincelar
Your answer

tirotear
Your answer

garabatear
Your answer

amartillar
Your answer

abanicar
Your answer

Never submit passwords through Google Forms.
Instructions

Each item in this form will prompt you with a word in the target language (Spanish). There are a total of twenty five (25) items. For each item, please provide the English translation of the word in the field below. If you do not know the translation of the word, leave the field blank.

engullir

Your answer

hachear

Your answer

enjabonar

Your answer

rastrillar

Your answer

campanear

Your answer
Instructions

Each item in this form will prompt you with a word in the target language (Spanish). There are a total of twenty five (25) items. For each item, please provide the English translation of the word in the field below. If you do not know the translation of the word, leave the field blank.

aplanar
Your answer

voltear
Your answer

espolvorear
Your answer

cepillar
Your answer

serruchar
Your answer
Instructions

Each item in this form will prompt you with a word in the target language (Spanish). There are a total of twenty five (25) items. For each item, please provide the English translation of the word in the field below. If you do not know the translation of the word, leave the field blank.

encerar
Your answer

tamborear
Your answer

machucar
Your answer

abucheear
Your answer

desenvainar
Your answer

BACK   SUBMIT

Never submit passwords through Google Forms.
B. Variables obtained from questionnaires

The following variables were used to assess participants through the studies presented in this thesis.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>MyAge</td>
<td>How old are you?</td>
</tr>
<tr>
<td>Gender</td>
<td>Gender: (male, female, prefer not to say, other)</td>
</tr>
<tr>
<td>VR Exposure</td>
<td>How much exposure do you have to Virtual Reality systems?</td>
</tr>
<tr>
<td>Languages Spoken</td>
<td>How many languages do you speak?</td>
</tr>
<tr>
<td>Languages Studied</td>
<td>How many languages have you studied in the past?</td>
</tr>
<tr>
<td>Language Difficulty</td>
<td>How difficult do you find learning a new language?</td>
</tr>
<tr>
<td>Intuitiveness</td>
<td>How intuitive did you find the learning experience to be?</td>
</tr>
<tr>
<td>MyBody</td>
<td>How much did you feel that the virtual body you saw when you looked down at yourself was your own body? (Despite it might not look like you)</td>
</tr>
<tr>
<td>Mirror</td>
<td>How much did you feel that the virtual body you saw when you looked at yourself in the mirror was your own body? (Despite it might not look like you)</td>
</tr>
<tr>
<td>LikeMe</td>
<td>How much did you feel that your virtual body resembled your own (real) body in terms of shape, skin tone, or other visual features?</td>
</tr>
<tr>
<td>Two Bodies</td>
<td>How much did you feel as if you had two bodies?</td>
</tr>
<tr>
<td>Perceived Age</td>
<td>Of what age was your virtual body?</td>
</tr>
<tr>
<td>Younger</td>
<td>How much younger did you feel?</td>
</tr>
<tr>
<td>Older</td>
<td>How much older did you feel?</td>
</tr>
<tr>
<td>Felt Age</td>
<td>What age did you feel yourself to be?</td>
</tr>
<tr>
<td>Felt Like</td>
<td>During the experiment, you felt like: (Child, Adult)</td>
</tr>
<tr>
<td>Child</td>
<td>How much did you feel like a child?</td>
</tr>
<tr>
<td>Engagement</td>
<td>How engaging did you find the technology enhanced learning experience?</td>
</tr>
<tr>
<td>Repeat</td>
<td>How inclined would you be to learn new vocabulary through this new virtual reality method?</td>
</tr>
<tr>
<td>Open Comment</td>
<td>Please write down any other comment or feeling you had regarding your experience.</td>
</tr>
</tbody>
</table>
C. Selected words for Embodied Learning Experiments

The following words and their translation in English were used for the user studies presented in Chapter 4 and Chapter 5.

<table>
<thead>
<tr>
<th>Spanish word</th>
<th>English Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atornillar</td>
<td>Screw</td>
</tr>
<tr>
<td>Embrujar</td>
<td>Jinx</td>
</tr>
<tr>
<td>Apuñalar</td>
<td>Stab</td>
</tr>
<tr>
<td>Rapar</td>
<td>Shave</td>
</tr>
<tr>
<td>Pincelar</td>
<td>Paint</td>
</tr>
<tr>
<td>Tirotear</td>
<td>Shoot</td>
</tr>
<tr>
<td>Garabatear</td>
<td>Doodle</td>
</tr>
<tr>
<td>Amartillar</td>
<td>Hammer</td>
</tr>
<tr>
<td>Abanizar</td>
<td>Fan</td>
</tr>
<tr>
<td>Engullir</td>
<td>Devour</td>
</tr>
<tr>
<td>Hachear</td>
<td>Axe</td>
</tr>
<tr>
<td>Enjabonar</td>
<td>Lather</td>
</tr>
<tr>
<td>Rastiriar</td>
<td>Rake</td>
</tr>
<tr>
<td>Campanear</td>
<td>Ring</td>
</tr>
<tr>
<td>Aplanar</td>
<td>Flatten</td>
</tr>
<tr>
<td>Voltear</td>
<td>Flip</td>
</tr>
<tr>
<td>Espolvorear</td>
<td>Sprinkle</td>
</tr>
<tr>
<td>Serruchar</td>
<td>Saw</td>
</tr>
<tr>
<td>Encerar</td>
<td>Wax</td>
</tr>
<tr>
<td>Tamborear</td>
<td>Drum</td>
</tr>
</tbody>
</table>
Bibliography


88


E. Langer, Counterclockwise, Hodder and Stoughton..


