Grounding body ownership and language in action: evidence from healthy and damaged brains

Klaudia Grechuta

TESI DOCTORAL UPF / ANY 2019

DIRECTOR DE LA TESI
Prof. Dr. Paul F.M.J. Verschure
Departament de Tecnologies de la Informació i les Comunicacions

Universitat Pompeu Fabra
Barcelona

To all that inspired this work and those who this work will inspire

Acknowledgements

My great gratitude goes to my supervisor Paul Verschure, for his vision and passion, for giving me the freedom I wanted and the guidance I needed. Many sincere thanks go to my students, colleagues, coauthors, and collaborators for inspirational discussions, insightful comments, encouragement, and a lot of fun. I truly thank my friends who have always been there for me and believed in me. I am also deeply thankful to my parents for being beautiful people, with young spirits, who inspired me on how to take life and showed me what really matters. A big piece of my heart goes to my sister, for being the best sister on Earth, and for our mission *Na koniec swiata i jeszcze dalej.* Last and foremost my love goes to Giovanni for inspiring me, believing in me, making me laugh, and for holding my hand throughout this journey, and many others, and many more...

Without you this here would not have been possible
...and it would not have been such great times.

Abstract

Contrary to the classical theory of mind, largely inspired by dualism, the theoretical framework of embodiment emphasizes the constitutive role of the body in the process of cognition. This view also referred to as the sensorimotor approach has been supported by a number of clinical and empirical studies suggesting that at the basis of both lower-level processes and higher-level functions is the coupling between the body, brain, and environment through action. Critically, these results challenge the traditional theories which view both the experience of the body and the conceptualization of the world through language as isolated and decoupled from action. These two abilities, in turn, seem functionally grounded in experience and essential to successfully act in the world by defining the boundaries of an embodied self, on the one hand, and enabling verbal communication, on the other. In an interdisciplinary effort which integrates methods such as bodily illusions and clinical trials, this dissertation comprises a series of experiments which aim at advancing our knowledge about the principles of body ownership and language (re)learning grounded in sensorimotor interactions of an individual with the world. In the first part of the dissertation, we present a set of four behavioral experiments that extend the classical multisensory theory of body ownership (i.e., based on the Rubber Hand Illusion). This approach indeed only accounts for externally-generated stimuli (i.e., tactile strokes) and neglects the integration of the efferent signals which are necessarily present during goal-oriented behavior. To bridge this gap, we asked whether body ownership is coupled to the motor system such that it depends on the congruency of sensory consequences of goal-oriented actions. Our results suggest that the mechanisms which underlie body ownership go beyond the mere integration of passively received multi-sensory signals and support the role of action, goals, and environment in building the minimal representation of the embodied self. In the second part of the dissertation, we tested the premises of the embodiment thesis in the context of language (re)learning. In particular, we designed a virtual reality-based contextualized multimodal therapy for the rehabilitation of language in post-stroke patients with nonfluent aphasia which capitalizes on the sensory-motor grounding of linguistic functions predicted by the empirical framework of embodiment. We tested, in a longitudinal clinical trial, the benefits of this approach compared to the standard therapy and showed that behaviorally relevant training of language in the context of action indeed promotes the recovery and retention of linguistic functions. Moreover, we demonstrated that the sensorimotor cueing embedded in our approach facilitates word retrieval. Altogether, the results of this dissertation contribute to the understanding of how both lower- and higher-level cognitive functions, such as body-ownership and language, are tightly coupled to the motor system, goals at hand, and the dynamics of the surrounding environment. These findings can contribute to the basic research as well as applied sciences with an emphasis on learning and rehabilitation.

Resum

Contràriament a la teoria clàssica de la ment fonamentalment inspirada pel dualisme, el marc teòric d'"embodiment"emfatitza el rol constitutiu del cos en els processos cognitius. Aquesta perspectiva, també referida com a aproximació sensoriomotora, ha rebut suport per part de nombrosos estudis empírics, suggerint que a la base dels processos cognitius tant de baix com d'alt nivell s'hi situa la interacció entre el cos, el cervell i l'entorn, mitjançant l'acció. Críticament, aquests resultats desafien les teories tradicionals sobre la mateixa experiència del cos com a procés de baix nivell, així com les aproximacions a l'adquisició del llenguatge com a procés d'alt nivell. Ambdues funcions, alhora, semblen essencials per, d'una banda, poder actuar al món satisfactòriament, definint el mínim sentit de "self", i d'altra banda, per habilitar la comunicació verbal. En un esforç interdisciplinari que integra mètodes tals com illusions corporals i assajos clínics aleatoritzats, aquesta dissertació es presenta com a una sèrie d'experiments amb l'objectiu de millorar el nostre coneixement sobre els principis de "body ownershipï del (re)aprenentatge del llenguatge basats en les interaccions sensoriomotores d'un invididu amb el món. Amb aquesta finalitat, en un conjunt de quatre experiments de comportament amb persones sanes, hem intentat estendre la teoria clàssica multisensorial de "body ownership", la qual es basa en paradigmes que no tenen en compte la integració de cap senyal eferent. Específicament, hem provat si el "body ownershipïnteracciona amb el sistema motor en el sentit de si depèn de senyals sensorials (incloent-hi les externes al cos) pertanyents a una tasca orientada a objectius. En un estudi posterior, vam posar a prova les nocions de connectivitat, espai peripersonal, i plausibilitat física com a condicions necessàries per "body ownership"en el context de l'acció. Finalment, ens vam preguntar quin és el rol del model intern que tenim de l'entorn en l'experiència del "self". Parallelament, inspirats per descobriments recents, vam posar a prova les premisses de la tesi d'"embodiment"en el context de (re)aprenentatge del llenguatge. En particular, vam dissenyar una teràpia comportamental, orientada a objectius, i multimodal, per pacients amb afàsia no fluent. Vam provar si aquesta aproximació és beneficiosa per a la

recuperació i retenció de les funcions de llenguatge, en comparació amb la teràpia estàndard basada en la visió proposicional. Conjuntament, els nostres resultats contribueixen a l’enteniment de com processos de baix nivell, com l’experiència que el cos ens pertany, i les funcions cognitives de baix nivell, com el (re)aprenentatge del llenguatge, estan íntimament lligats al sistema motor, els objectius, i les dinàmiques de l’entorn. Més enllà de la seva rellevància per a la recerca en neurociència, aquests resultats poden tenir aplicacions en l’àmbit de la rehabilitació de desordres del bodily-self i del llenguatge.

Publications

Included in the thesis

Peer-reviewed

Grechuta, K., Ballester, B. R., Munné, R. E., Bernal, T. U., Hervás, B. M., Mohr, B., Pulvermuller, F., San Segundo, R., & Verschure, P. (2019) Automated multisensory cueing facilitates recovery in aphasia: a triple-blind clinical study. *Journal of NeuroEngineering and Rehabilitation.* (Under review)

Grechuta, K., De la Torre, J., Rubio Ballester, B., & Verschure, P. (2019) Challenging the boundaries of the physical self: purely distal cues in the environment impact body ownership *Royal Society Open Science.* (Under review)

Grechuta, K., Ballester, B. R., Munné, R. E., Bernal, T. U., Hervás, B. M., Mohr, B., Pulvermuller, F., San Segundo, R., & Verschure, P. (2019) Augmented embodied dyadic therapy boosts recovery in patients with nonfluent aphasia: a randomised controlled trial. *Stroke.*

Grechuta, K., Ulysse, L., Rubio Ballester, B., & Verschure, P. (2019). Self beyond the body: action-driven and task-relevant purely distal cues modulate performance and body ownership. *Frontiers in Human Neuroscience.*

Grechuta, K., Ballester, B. R., Munné, R. E., Bernal, T. U., Hervás, B. M., San Segundo, R., & Verschure, P. F. (2017, July). The effects of silent visuomotor cueing on word retrieval in Broca's aphasies: A pilot study. In *International Conference on Rehabilitation Robotics 2017.*

Grechuta, K., Guga, J., Maffei, G., Ballester, B. R., & Verschure, P. F. (2017). Visuotactile integration modulates motor performance in a perceptual decision-making task. *Scientific Reports.*

Grechuta, K., Ballester, B. R., Duff, A., Oller, E. D., Pulvermuller, F., & Verschure, P. F. (2016). Intensive language-action therapy in virtual reality for a rehabilitation gaming system. *Journal of Pain Management.*

In preparation

De la Torre, J., Grechuta, K., Daversa, D., & Verschure, P. (2019) (Body) ownership as a matter of control, not priors, connectedness or peripersonal space.

Other publications and abstracts

Peer-reviewed

Grechuta, K., Vouloutsi, V., & Verschure, P. (July, 2019) Towards psychologically plausible robots: evaluation of the iCub's facial expressions. In *Conference on Biomimetic and Biohybrid Systems.* (Under review).

Grechuta, K., Ballester, B. R., Munné, R. E., Bernal, T. U., Hervás, B. M., San Segundo, R., & Verschure, P. F. (2017, October). The effects of semantic auditory cueing on lexical retrieval in chronic Broca's aphasics. In *European Congress of NeuroRehabilitation 2017.*

Reidsma, D., Charisi, V., Davison, D., Wijnen, F., van der Meij, J., et al. (2016, July). The EASEL project: towards educational human-robot symbiotic interaction. In *Conference on Biomimetic and Biohybrid Systems.*

Blancas, M., Vouloutsi, V., Grechuta, K., & Verschure, P. (2015, July). Effects of the robot's role on human-robot interaction in an educational scenario. In *Conference on Biomimetic and Biohybrid Systems.*

Zegarek, G., Dalmazzo, D., Brandi, S., Grechuta, K., Arsiwalla, X., & Verschure, P. F. (2015) Mapping the Aphasia Connectome with BrainX3:

Connectomic Predictions of TMS-Defined Eloquent Cortex. In *Neuroscience R&D Technologies Conference.*

Lallee, S., Vouloutsi, V., Munoz, M. B., Grechuta, K., Llobet, J. Y. P., Sarda, M., & Verschure, P. F. (2015). Towards the synthetic self: making others perceive me as an other. *Paladyn, Journal of Behavioral Robotics.*

Vouloutsi, V., Munoz, M. B., Grechuta, K., Lallée, S., Duff, A., Llobet, J. Y. P., & Verschure, P. F. M. J. (2015). A new biomimetic approach towards educational robotics: the distributed adaptive control of a synthetic tutor assistant. In *New Frontiers in Human-Robot Interaction.*

Gou, M. S., Vouloutsi, V., Grechuta, K., Lallée, S., & Verschure, P. F. (2014, July). Empathy in humanoid robots. In *Conference on Biomimetic and Biohybrid Systems.*

Bampatzia, S., Vouloutsi, V., Grechuta, K., Lallée, S., & Verschure, P. F. (2014, July). Effects of gaze synchronization in human-robot interaction. In *Conference on Biomimetic and Biohybrid Systems.*

Vouloutsi, V., Grechuta, K., Lallée, S., & Verschure, P. F. (2014, July). The influence of behavioral complexity on robot perception. In *Conference on Biomimetic and Biohybrid Systems.*

In preparation

Herreros I., Miquel L., Nuno L., Rubio Ballester B., Grechuta K., Gual T., Balcells-Oliveró M., & Verschure P. (2019) Impaired implicit motor adaptation in chronic cannabis users

Ballester, B. R., Grechuta, K., & Verschure, P. F. (2019) Closing the Sensorimotor loop.

Contents

List of Figures

List of Tables

Chapter 1

INTRODUCTION

"Who in the world am I? Ah, that's the great puzzle."

Alice in Wonderland

The seemingly simple question which fascinates Alice and which has inspired the work presented in this dissertation constitutes one of the most fundamental ontological questions that has intrigued philosophers and scientists for thousands of years and that, despite numerous theoretical and empirical efforts aiming to resolve it, still remains elusive. Indeed, if we carefully analyze Alice's question of "Who in the world am I?," we can immediately notice that in order to answer this question we need to define two main elements which it is comprised of: the *I*, on the one hand, and the *world*, on the other. Classical philosophy and modern cognitive science have started from the assumption that the *I* and the *world* can be, by nature, functionally independent and thus they studied these concepts as separate and in isolation from one another.

For instance, Plato, one of the fathers of the dualist thought, considered the physical world to be a rough copy of the real world (or the world of the ideas) and advanced a clear division between the matter and the soul [Murphy, 2006]. The world of ideas, formed of true and pure concepts, could exclusively be experienced by the mind through the power of intellect. On the contrary, the body constituted a part of the material world, and it could only access the physical instantiation of the ideas, which were considered fallacious and misleading [Bloom and Kirsch, 2016]. Crucially, the idea that the body equipped multiple senses (i.e., the *I*) cannot access the real world of forms makes the two become essentially decoupled, the notion which is central to Cartesian philosophy. Descartes believed that all experiences and knowledge acquired from and through the senses could not be trusted [Wertheimer, 2012]. In his *Meditations On First Philosophy* [Simmons, 2003], Descartes doubted the existence of shapes, movements, and even the body itself, reaching the radical conclusion that the only truth one can be certain of is the act of thinking. Thus here again the physical body is considered a pure material substance, which automatically operates in the world but is functionally detached from the mind that underlies thinking and reasoning.

The dualist view has largely inspired more recent theories, such as the computationalism [Putnam, 1963], which played a central role in the development of modern cognitive science and psychology during the 1960s and 1970s and which is still one of the most influential theories in the contemporary philosophy of mind [Chalmers, 1993]. The central assumption of the computational theory of mind, initially proposed by Hilary Putnam, is that the mind acts like an information processing system and that all cognitive processes (from perception and decision-making to language) stem from computations physically implemented by neurons in the brain [Putnam, 1963, Fodor, 1975]. Critically, in this view, the mind is interpreted as a "software", or a Turing machine, whose role is to generate series of outputs by manipulating sequences of inputs according to logical and mathematical rules. To interface with the physical world through senses, the abstract mind needs to convert percepts into symbols through a process of (mental) representation. Those, in turn, contain information about the input which can be interpreted. Symbols refer to their referents in an abstract and perhaps arbitrary way, and they are essentially decoupled from the physical world, similar to the platonic world of ideas. In this sense, this theory is closely related to the representational theory of mind as well as the language of thought [Fodor, 1975]. The critical implication of these theories, however, is that they abstract from particular physical properties of the “software” or the mind, which in principle means that the body is decoupled from all the computational processes.

Despite great contributions to the development of highly intelligent programs, which can play Chess and take part in conversations, computational theories of mind have been largely challenged on the notion of “understanding”. Indeed, describing the mind as a symbolic processor incurs into a conceptual paradox, as demonstrated in the Chinese Room argument [Searle, 1980]. Among others [Chalmers, 1993, Van Gelder and Port, 1995], the Chinese Room argument has gained significant attention and become one of the best-known arguments in recent philosophy 1.3) [Searle, 1980]. In his thought experiment, John Searle demonstrates that a non-Chinese speaker might indeed achieve excellent performance in answering simple questions by solely manipulating Chinese symbols according to strict rules (i.e., syntax). This does not mean, however, that the subject has a true understanding of what those symbols mean. In other words, this experiment shows that syntax alone is not sufficient to convey the semantic meaning of concepts and in this way, computational theories seem incomplete in explaining how can the mind links the symbols to their referents. Treating the mind as an isolated process, detached from the physical world, makes it therefore impossible to even define the world itself, leaving Alice’s question unanswered.

One could speculate that the failure of the approaches mentioned above is akin to their fundamental assumption that the mind is isolated and detached from its context [Chalmers, 1993, Van Gelder and Port, 1995, Searle, 1980, Glenberg and Robertson, 2000]. Interestingly, however, if we have a closer look at the opening question of “Who in the world am I?” and consider the preposition of place *in*, it implicitly suggests that the *I* is immersed and present in their surroundings. Moreover, it stresses the situatedness and the coupling of an individual with the world. This apparently negligible detail might, in fact, play a critical role in both approaching the question and interpreting the answers. Intending to shed light on Alice’s doubts and contribute to the understanding of how the brain builds the sense of self and meaningful representation of the world, we propose to look at an individual and the world, not as separate entities, but as strictly coupled and inseparable. In particular, we hypothesize that the cognitive processing of an individual, from the lower- to higher-level functions, can only be understood by looking at a system composed by an agent and its environment coupled by sensations and actions [O'Regan and Noë, 2001, Noë, 2004, Barsalou et al., 2003]. This premise is the basis of the theory of embodied cognition, also referred to as the sensorimotor approach, which has gained compelling recognition during the last two decades [Barsalou et al., 2003, Lakoff and Johnson, 1999, Clark, 1999, Wilson and Foglia, 2011, Seth, 2013, Verschure et al., 2003].

Indeed, the theories of embodiment have recently challenged the dominant hypotheses of the Cartesian dualism or computationalism, which regard the bodily mechanisms of sensory processing and motor control, as somewhat secondary or tangential to cognition. In particular, the thesis of embodiment proposes that all cognitive processes are *grounded*, that is, they are functionally coupled to the physical properties of the body including the sensory and motoric systems [Wilson and Foglia, 2011, Barsalou, 2008, Verschure et al., 2003, Pulvermüller, 2005, Gallagher, 2000]. Those properties play a constitutive or causal role within the cognitive process such that cognitive functions are acquired through and shaped by sensorimotor interactions of an embodied goal-oriented individual with their surrounding environment. Theoretical framework of embodied cognition asserts that these cognitive functions include both lower-level processes and higher-level functions, which is supported by an ample empirical evidence [Warrington and McCarthy, 1987, Lakoff and Johnson, 1999, Martin and Chao, 2001, Barsalou et al., 2003, Pulvermüller, 2005, Gallese and Lakoff, 2005, Kiefer and Pulvermüller, 2012].

How could we then tackle the question of Alice from the perspective of embodied cognition? On the one hand, a prerequisite to building a representation of self and understanding the *I* is to study the physical body. In essence, it is the body which allows defining the boundaries between the self and the world enabling interaction and situatedness [Blanke, 2012]. On the other hand, a prerequisite to understanding the *world* is to study how concepts are grounded in sensorimotor interactions of an embodied individual with the world, and how these concepts are represented through language. Thus, in the context of this dissertation, we will we will try to elucidate the two dimensions of Alice's question by studying the principles of body ownership (Part I) and language (Part II). To this end, we will use an interdisciplinary approach which integrates experimental behavioral studies employing bodily illusion and clinical trials.

Before starting the individual chapters (2-9), we provide a brief introduction to the fundamental concepts which allowed the design of the empirical work. We will contextualize them within the traditional framework as well as the current approach, that is, embodied cognition. Finally, we will provide an outline of the dissertation which summarizes each contribution and points out the relevant links between each chapter.

1.1 Embodied Cognition and Body Ownership

Different from classical thinkers and philosophers, who considered the body as a *nest* for the soul and psyche functionally independent from higher-level cognitive functions, contemporary psychology, and cognitive neuroscience has drawn much attention to the peculiarities of the physical self and the study of its biological, social, and phenomenological dimensions [Gallagher, 2000]. The notion of bodily awareness, in particular, has gained considerable interest being at crossroads of philosophical and scientific debates and research activities regarding the self, perception, action, and space. Within this framework, *body ownership* which underlies subjectivity and self-recognition seems to play a critical role in both lower-level processes and higher-level cognitive functions thus being one of the key features of self-awareness. In the following sections, we will outline the definition of body ownership, its functions, as well as the classical and more recent approaches to studying its principles. Finally, we will present several intriguing questions related to the experience of body ownership in the context of action, which are inspired by the thesis of the embodied cognition and which still require the support of empirical evidence. We will close this section by presenting how were these questions addressed in the context of the present thesis.

Defining body ownership

One rarely makes an error when asked to recognize their hand, localize it in space, and determine to whom it belongs. *Body ownership* is one of the fundamental conditions enabling minimal phenomenal selfhood, which, in general terms, refers to the experience of the body as belonging to oneself and being the source of sensations [Gallagher, 2000, Gallagher, 2008, Tsakiris, 2010, Limanowski and Blankenburg, 2013]. It allows us to maintain a stable representation of the physical self and accurately determine its boundaries, which are essential prerequisites to achieve goals, anticipate undesirable events, and thus successfully act in the world. Interestingly, the process of attributing a body to the self is often taken for granted, and many of the mechanisms which drive it, especially in action-contexts, still remain elusive.

What does body ownership mean in empirical terms? It has been proposed that body ownership depends on two cognitive indicators, that is, the *feeling* and the *judgment* [De Vignemont, 2014]. The former refers to the multisensory perceptual representation of the body which is mediated by top-down influences [Botvinick, 2004]. In particular, when the incoming signals from multiple sensory channels (i.e., visual, tactile, vestibular) are spatiotemporally aligned, this results in a sensation of ownership of a coherent body. Damasio [Damasio and Macaluso, 2004] described this state of the wholeness of the body as the so-called "background feeling" which is always present, even if it is not continuously consciously accessed or noticeable. The judgment of ownership, on the other hand, is conceptual in nature indicating the awareness of being the owner of a body.

Although under normal circumstances, the first-order feeling of ownership underlies the second-order judgment, somewhat yielding that they are functionally coupled, clinical studies support the notion that the two can be qualitatively dissociated. For instance, there have been neurological cases in which (1) patients judged that their hand belonged to them despite feeling it as 'alien', [Sacks and Sacks, 1998, Cole and Cole, 1995], (2) they neither felt nor judged that the limb belongs to them, and thus considered it as alien [Feinberg, 2002], or (3) they judged that a limb does not belong to them, despite feeling it like their own [Vallar and Ronchi, 2009]. These examples support that ownership can be attributed through both the *feeling* and the *judgment* and, therefore, for the scope of this dissertation, we will investigate the principles of body ownership focusing on both the experiential and rational indicators which will be reflected in the measures.

How to study the sense of body ownership?

In spite of compelling evidence about its essential role in action, perception, and cognition, the study of bodily awareness and in particular body ownership has received relatively less attention compared to other cognitive functions such as, for instance, visual perception. First, this might be due to the complexity of the phenomenon which requires an understanding of both uni- and multisensory processing of interoceptive (i.e., within the body) and exteroceptive (i.e., outside of the body) signals which drive both the physical and emotional states (e.g., thirst, sexual arousal). Secondly, this could be related to the very nature of body ownership which cannot be voluntarily controlled and, by nature, is always present: “the same old body always there” as William James summarized [James, 2013]. On a similar vein, Merleau-Ponty noted that “ [...] the permanence of my own body is entirely different in kind ... Its permanence is not a permanence in the world, but a permanence on my part” [Merleau-Ponty, 1945].

From the experimental perspective, being always present or, in other words, not having at least two states (i.e. “on,” “off”), constitutes a significant methodological drawback and pinpoints the final issue which relates to the robustness of the quantification of the experience of ownership. In fact, still today, the most common instrument to measure body ownership is a self-report which is a subjective method and therefore difficult to be reliably gauged. Only at times, the questionnaire is accompanied by behavioral or physiological tools [Botvinick and Cohen, 1998, Armel and Ramachandran, 2003] which will be discussed more in detail in the following sections. Together, investigating the principles of body ownership presents several methodological challenges and is, by nature, different from studying other types of awareness, for example, visual or tactile, which can be easily experimentally manipulated.

Bodily illusions as tools to study body ownership

How can we then study the sense of body ownership? During the last two decades, perceptual bodily illusions have become major tools in studying the basis of bodily self-attribution and self-location with healthy participants, providing insights to both the second-order judgment and the first-order experience [Gallagher, 2011]. As Smith pointed out “It is tempting to think that because our bodies are, as it were, so close to us, the scope for illusions here is minimal. In fact, however, recent research in this area has presented some of the most striking illusions in all the literature” [Smith, 2004].

The cornerstone in the empirical study of body ownership is the so-called Rubber Hand Illusion (RHI) [Botvinick and Cohen, 1998] where the ownership of a fake limb is induced via synchronous stroking of the real hand occluded to vision and a rubber hand placed in front of the participant. In the following section, we will discuss in more detail the classical approach to studying body ownership using RHI and the current hypothesis about its basis.

1.1.1 Body Ownership as a Sensory State

The Rubber Hand Illusion paradigm

A great inspiration for the empirical investigation of the principles of the self-attribution of the body was the Rubber Hand Illusion [Botvinick and Cohen, 1998], which is by now the most established bodily-illusion paradigm. RHI dates back to 1998 when Botvinick & Cohen [Botvinick and Cohen, 1998] observed that when participants received synchronous tactile strokes on their real hand (hidden from vision) and on an ipsilateral fake rubber hand (placed in front of them), they would report that the feeling of the perceived touches occurred on the rubber hand and not their own (Figure 1.1 A). More interestingly, the subjects further reported a feeling as if the fake rubber hand belonged to their body. Indeed, when asked to indicate the location of the real hand, they would make pointing errors exhibiting a significant bias towards the location of the rubber hand, the so-called proprioceptive drift.

The attribution of the rubber hand to the representation of the body and the first-order feeling of ownership was supported by subsequent studies which used physiological measures to quantify ownership. For instance, threatening the rubber hand in the condition of congruent stroking resulted in increased arousal, measured through galvanic skin response, as compared to the incongruent condition [Armel and Ramachandran, 2003, Ehrsson et al., 2007, Gentile et al., 2013]. Other physiological indicators of the illusion included increased histamine reactivity on the real hand [Barnsley et al., 2011], which signals a down-regulation of the immune system, as well as a drop in skin temperature of the real hand [Moseley et al., 2008].

Together, this evidence suggests that during RHI the increased self-attribution of the rubber hand is accompanied with a decreased self-attribution of the real hand or, in other words, the rubber hand temporarily replaces the real one [Longo et al., 2008].

Necessary and sufficient conditions for body ownership

Initially, the results of the classical Rubber Hand Illusion [Botvinick and Cohen, 1998] and its numerous replications yielded that a necessary condition for the emergence of ownership is the congruency of the multisensory cues. Indeed, none of the above effects takes place in the control condition in which the visual and tactile cues are either temporarily or spatially asynchronous [Costantini and Haggard, 2007]. Thus, first interpretations of the processes underlying body ownership held that it strictly requires perceptual, in this case, visual and tactile, correlations. This proposal was later formalized as the *multisensory hypothesis of body ownership* [Ehrsson, 2012] which posited that body ownership is a passive sensory state at the basis of which is the integration and matching of information from mul-

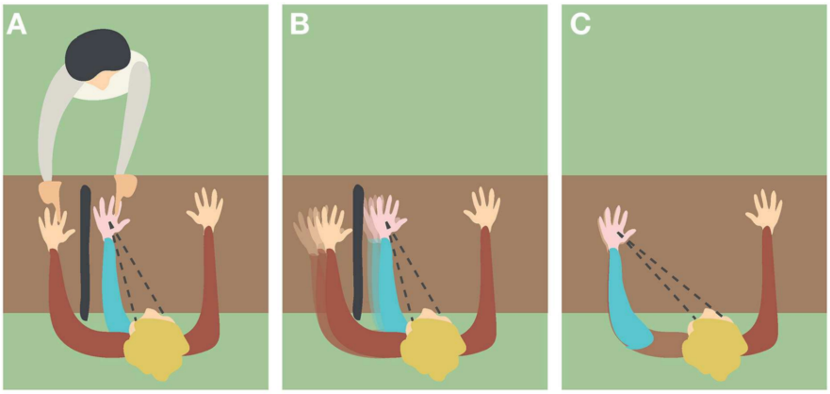


Figure 1.1: (A) Visuotactile integration. The participant is observing a fake rubber hand placed in a plausible position in from of them being stroked while receiving congruent tactile stimulation on the real hand occluded to vision. (B) Visuomotor integration. The participant is performing arm movements with his/her arm occluded to vision while observing a fake (i.e., rubber) or virtual hand which moving synchronously. (C) Visuoproprioceptive integration. The participant is observing a fake hand place in a plausible position with his/her real hand occluded to vision (from [Kilteni et al., 2015]).

tiple sensory channels [Maravita et al., 2003, Makin et al., 2008, Ehrsson, 2012, Tsakiris, 2010, Blanke, 2012].

Multisensory integration occurs when two or more sensory signals come from approximately the same location (the 'spatial rule') and at approximately the same time (the 'temporal rule') [Holmes and Spence, 2005, Stein and Stanford, 2008]. Under these conditions, it is likely to interpret, feel, and judge that the two (or more) signals have a common source which pertains information about it and, in the particular case of the RHI, it *is* the rubber hand. According to this hypothesis, multisensory integration was both necessary and sufficient for self-attribution of the body exhibited through first-order experience and second-order judgment of its ownership.

Interestingly, subsequent experiments have shown that the illusion of ownership in the context of RHI is significantly modulated by top-down influences originating from the prior knowledge about the body driven by the experience. Hence, it suggested that the visuotactile correlations driving the illusion are indeed necessary, but not sufficient [Tsakiris and Haggard, 2005]. For instance, a number of experiments [Tsakiris and Haggard, 2005, Costantini and Haggard, 2007, Lloyd, 2007, Makin et al., 2008] demonstrated that there seems to be an active process which allows comparison between the viewed object (i.e., rubber hand) and the reference model of the own body acquired through experience. This (internal) model contains information about visual, anatomical and structural properties of the body such that through the top-down comparison process, ownership cannot be transferred to, for instance, non-corporeal objects or those located in an implausible position. Conversely, RHI depends on the internal body representation and strictly requires physical, anatomical, postural and spatial plausibility of the real and fake hands within the peripersonal space [Tsakiris and Haggard, 2005, Costantini and Haggard, 2007, Lloyd, 2007, Makin et al., 2008] (see also [Liepelt et al., 2017, Armel and Ramachandran, 2003]).

Neural correlates of the Rubber Hand Illusion

Using functional Magnetic Resonance Imaging (fMRI) and Positron Emission Tomography (PET), several neuroimaging studies investigated the neural signatures of body ownership during the rubber hand illusion. Results revealed that the activity in the frontal and parietal areas was associated with the very experience of ownership. Specifically, some studies reported increased BOLD (Blood Oxygen Level-Dependent imaging) responses in the ventral premotor cortex in both hemispheres and intraparietal sulcus during the synchronous stroking of the rubber hand placed in an anatomically plausible position [Ehrsson et al., 2004]. These results were replicated by some [Ehrsson et al., 2007, Gentile et al., 2013] but not all [Tsakiris et al., 2007] subsequent studies. Another imaging studies less consistently reported the involvement of regions such as somatosensory cortex [Tsakiris et al., 2007, Makin et al., 2008], the right posterior insula [Tsakiris et al., 2007], and the temporoparietal junction [Tsakiris, 2010] at different stages of the RHI. Overall, these results converge to suggest that the illusion of ownership is associated with increased activity in regions involved in the processing of multisensory signals. However, the consensus is still missing on the exact contributions of each cortical region to each phase of the illusion.

A recent study investigated electrophysiological correlates of body ownership using RHI with five patients implanted with arrays of subdural electrodes which measured cortical surface potentials [Guterstam et al., 2018]. The results yielded that the feeling of ownership of the rubber hand was reflected in increased activity in neuronal populations of the premotor and intraparietal cortices thus providing a link between the electrical activity that is recorded at the cortical surface and previously discussed BOLD responses. The tactile stimulation (i.e., stroking) modulated intraparietal activity more than the premotor cortex supporting that these two key areas might play different roles in the process of self-attribution of the artificial limb. More studies are still needed to provide a unified view on the cortical basis for the shaping and the feeling of limb ownership.

Does the Rubber Hand Illusion explain it all?

In the previous sections, we have shown that in the context of the classical RHI, in which individuals are *passively* receiving spatiotemporally aligned tactile strokes on their own hand (occluded to vision) and a fake hand (placed in front of them), body ownership relies on two necessary processes. These processes include (1) bottom-up accumulation and integration of multisensory cues (i.e., tactile and visual), and (2) top-down comparison between the novel sensory stimuli (i.e., artificial rubber hand) and experience-driven priors about the internal model of the body [Tsakiris and Haggard, 2005, Blanke, 2012, Clark, 2013, Seth, 2013, Apps and Tsakiris, 2014]. At the neural level, attribution of an artificial hand into the body representation is associated with increased activity in a range of multisensory regions which seem to underlie different phases of this process [Tsakiris et al., 2007,Makin et al., 2008,Gentile et al., 2013,Guterstam et al., 2018].

Curiously, the classical approach to studying ownership using the rubber hand illusion solely sheds light on the mechanisms underlying body ownership when the individuals are passively exposed to sensory (i.e., tactile) cues, while the principles of ownership in the context of self-generated action remain largely unexplained. In fact, RHI purposely isolates the pure experience of body ownership in a complete absence of movement and any efferent sensory information [Tsakiris et al., 2006a]. Thus although from the experimental perspective, RHI seems to offer a useful mean of manipulating body ownership and provides insights on its behavioral and neural underpinnings during passive exposure to tactile cues, this paradigm seems incomplete in at least two ways.

First, rubber hand illusion seems to assume that there exists a sensory state in which an organism can only passively receive sensory stimuli, almost like a brain placed in a vat and connected to external input sources [Dennett, 2017]. Such approach to studying perception, and in particular bodily self-awareness, in isolation from action seems at odds with the proposal that the brain has evolved to sense and act with respect to the individual's (dynamical) internal and external states [James, 1890]. Besides, it violates a more general principle of embodied cognition which views the body as the mean to physically engage with the world through goal-oriented action [Merleau-Ponty, 1945, Verschure et al., 2003]. The absence of voluntary movement further pinpoints a second limitation of the paradigm, namely, that the current model of the rubber hand illusion [Tsakiris, 2010, Blanke, 2012] indeed fails to explain the principles of body ownership when the sensory signals are self-generated through behavior. This seems to be a critical aspect since it has been demonstrated that when we move the top-down comparison between the novel stimuli and the priors about the body (which is a necessary process for the experience of RHI) is not required for the feeling of ownership to emerge. In particular, recent studies suggest that in action-contexts we might self-attribute any effector which moves in accordance with our motor commands, including a robotic body, virtual objects or even a disconnected balloon [Banakou et al., 2013, Peck et al., 2013, Romano et al., 2015, Van Dam and Stephens, 2018, Ma and Hommel, 2015a].

With the goal of bridging theoretical and empirical gaps between the classical hypothesis of body ownership, based on RHI, and the current embodied perspective in the next section we will present recent sensorimotor approaches to studying body ownership. Later, we will outline still unsolved intriguing questions about the principles of body ownership in the context of goal-oriented action, and briefly explain how these were addressed within the scope of this thesis.

1.1.2 Body Ownership in the Context of Action

Embodied cognition perspective

As presented in the “Phenomenology of Perception” by Merleau-Ponty [Merleau-Ponty, 1945], the thesis of embodiment differs largely from the classical theories as to the function of the body in the process of cognition. In particular, the main premises include: (1) the body is not an object which can be represented but rather a subject of cognition, (2) the presence of the body refers to the presence of the body in a dynamical world, and (3) the body which we experience is the body which *acts*. Consequently, the followers of the embodied view talk about the so-called *lived body*, or the *lived body-environment*, which is understood in terms of goal-oriented sensorimotor engagements of an individual with the world through action [Gallagher, 1986]. In this context, the lived body assumes the readiness to act, or in other words “I can”, and the bodily space indicates the action-space which allows the action and which is determined by the objects in the environment, their affordances, subjective skills, and the current goal/s. As such, this space is flexible, and it can change dynamically depending on the current state of an agent and the world.

This view, which explicitly posits that action and in particular interaction of the body and the world sits at the core of cognition gave rise to a new sensorimotor approach [Zahavi and Parnas, 1998, O’Regan and Noë, 2001, Gallagher, 2006]. According to the premises of this framework, introduced by early embodiment [Merleau-Ponty, 1945, Gallagher, 2000] and later cognitive neuroscience [Pecher and Zwaan, 2005, Gottlieb, 2007, Verschure et al., 2003, O'Regan and Noë, 2001, O'Regan, 2011], the dynamical body is in continuous bidirectional exchange with its dynamical environment through action. This view differs fundamentally from the empirical scope of the Rubber Hand Illusion (RHI) where the main goal is to understand the principles of body ownership in isolation from voluntary action and, consequently, independent from any reafferent information [Botvinick and Cohen, 1998, Tsakiris, 2010, Blanke, 2012]. Conversely, embodied cognition argues in favor of the constitutive role of action in lower level processes and higher level functions which supports that at the basis of the experience of ownership is the combination of reafferent and exafferent information from multiple sensory channels.

Neurophysiological evidence in favor of the embodied theory of body ownership: the case of tools

A growing body of neurophysiological evidence suggests that the representation of the body might be altered at the neural level when, for instance, using a tool [Maravita and Iriki, 2004]. In particular, studies demonstrate enlarged visual Receptive Fields (RF) of bimodal neurons coding for the hand peripersonal space in monkey's parietal cortex after training with a rake [Iriki et al., 1996]. In other words, tactile RFs on the hand gradually displace their visual receptive field from an initial position close to the hand towards the tip of the tool, as a function of learning to use this same tool. Tool-use studies involving neurological patients with unilateral lesions and healthy human subjects also show changes in multisensory interactions in the peripersonal space [Maravita and Iriki, 2004, Serino et al., 2015]. In the former, active use of a tool seems to improve the cross-modal links between tactile stimuli occurring on the hand and the visual stimuli on the tip of the tool [Farnè and Làdavas, 2000]. Such neural alterations might underlie the historical observations that an individual using a tool experiences tactile sensations *at the tip of the tool* [Descartes, 1958]. In fact, when asked to cross two sticks to which mechanical stimulation is delivered, subjects experience difficulties in judging the temporal order of the stimulation, the same as if it was delivered to one's crossed hands [Yamamoto and Kitazawa, 2001]. Thus, the physical-self seems to be located within an action space which incorporates the tool as proposed by the embodied hypothesis [Merleau-Ponty, 1945, Gallagher, 1986].

The studies discussed above suggest that, in principle, in the context of goal-oriented action (Figure 1.1 B), body ownership neither depends on the prior knowledge about the properties of the physical self nor a fixed peripersonal space, as reported in the RHI [Tsakiris and Haggard, 2005, Blanke, 2012, Clark, 2013, Seth, 2013, Apps and Tsakiris, 2014], but rather on all the sensory cues which pertain to the action and to the goal at hand. Indeed, this mechanism seems to be similar to the one of the internal forward model or corollary discharge which guides motor behavior [Wolpert and Flanagan, 2001, Miall and Wolpert, 1996, Crapse and Sommer, 2008]. Specifically, the role of a forward model is to predict the sensory consequences of the motor system of an individual in response to self-initiated movement. This allows the central nervous system to estimate the current and future state of the effector immediately and without peripheral information. On the one hand, the actual state might be then compared to the intention which drives corrective motor behavior and, on the other, it might be informative of body ownership. For example, if I move my hand to the right and I see that the hand is moving to the right, likely it is my hand. Conversely, if I move my hand to the right and I see that the hand is moving to the left, likely it is not my hand. This has been reported in some neurological cases also referred to as the “alien hand” [Vallar and Ronchi, 2009]. Curiously, however, this mechanisms has not been explicitly addressed by the yet few studies which used a physical or virtual reality adaptation of the so-called *moving* RHI [Dummer et al., 2009, Walsh et al., 2011, Tsakiris et al., 2006b, Kammers et al., 2009, Newport et al., 2010, Sanchez-Vives et al., 2010, Kalckert and Ehrsson, 2012, Shibuya et al., 2018].

Grounded in the theory of the embodiment and recent cognitive science, which support the notion that self-awareness is inseparable from bodily activities and sensorimotor expectations [O’Regan and Noë, 2001, Noë, 2004, Verschure et al., 2003], in the present dissertation we aimed at addressing some of the critical questions which regard the principles of body ownership when acting.

Body ownership in action: open questions

Grounded in the hypothesis of embodiment, and its emphasis of the constitutive role of the body in cognitive processes, the first topic of this dissertation (Part I) presents a set of studies investigating different principles of body ownership in action. This work is comprised of four questions reflected in four different studies with healthy subjects.

First, the coupling of the body and environment through action suggests that there might be a functional relationship between the experience of ownership and motor control. We shed light on this question in the first study (Part I). Secondly, if as Merleau-Ponty proposes, the body is a *vehicle* allowing us to act in a goal-oriented way, then body ownership might depend on all the sensory cues which pertain to this action and the goal, which is the question that we address in the second study (Part I). Third, we aimed at understanding to which extent does this *vehicle* need to be bodily in physical terms. If my effector does not have a bodily shape and is away from my peripersonal space but I *can* control it and it *is* a tool to achieve my goals, can I experience it as part of me? This was the question of the third study (Part I). Finally, what is the role of the world in the ownership of my body? I do not only predict the consequences of my actions, but also those of the events of the external environment. For example, I know that if I throw confetti, it will immediately fall on the floor. The violation of motor predictions was shown to affect the ownership of my body. Can a violation of the predictions about the events in the world also affect my ownership? From the embodied perspective, in which body is tightly coupled to the external world, one could speculate that every sensory signal matters for a robust precept including the precept of self. This question has been experimentally tested, and the results are presented in the fourth and the last study related to body ownership in action.

1.2 Embodied Cognition and Language

What is the role of the body in higher-level cognition? Merleau-Ponty views embodiment in a double sense such that, on the one hand, the body is an experiential structure and, on the other, it is a context or milieu of higher-level cognitive functions [De Beauvoir, 1945]. Thus, understanding the principles of self-awareness including body ownership in the context of action, addressed in Part I of this dissertation, seems at the core of understanding the role embodiment in lower-level processes. The next step and the next goal of this work are to prompt the thesis of embodiment by investigating how cognition is dependent on the body. Does the embodiment play a significant causal or physically constitutive role in higher-level cognitive functions such as language processing [Clark, 1999, Verschure et al., 2003, Gallagher, 2006, Lakoff and Johnson, 2008]? This question pertains to the seconds topic of this dissertation in which we study whether the principles of the embodiment can be used to repair functions of damaged brains. This work is introduced in the following sections and systematically presented in the context of clinical work in Part II of this dissertation.

1.2.1 Classical View of Concepts

Dating back to the ancient Greek philosophers, the debate regarding the nature of conceptual and linguistic representations is more than 2000 years old [Runes, 1960]. It is already during the times of Plato that theorists and philosophers questioned the origin of concepts and speculated whether they are an innate quality of human mind or they derive from sensory experiences triggered by interactions with the environment. Interestingly, Plato himself, and later, other modern philosophers including Rene Descartes, Gottfried Wilhelm Leibniz, and Immanuel Kant did not support the hypothesis that concepts arise from perception and action [Markie, 2004]. Instead, embracing the rationalist assumptions, they would regard concepts as well as ideas as either innate and a priori or as conceived through the process of reasoning. Within such premise, concepts were classified as fundamentally different from sensory experiences which, in turn, were viewed as rather unstructured, misleading and, therefore, not well suited as the basis for meaning and knowledge.

Comparable debate on the role of sensory and motor representations in the construal of meaning has been present in contemporary philosophy and early cognitive science influenced by predicate logic as well as propositional and computational formalisms [Fodor, 1983, Pylyshyn, 1984]. As discussed in the previous sections, modern cognitive psychology traditionally regarded the body as peripheral to cognition. Similar, the fundamental bodily properties, that is, the ability to sense and the ability to act were considered as secondary low-level processes. In particular, they were seen as mutually independent (i.e., perception did not affect action and vice versa) and irrelevant to high-level cognitive functions such as conceptual and language processing [Pylyshyn, 1999]. Instead, symbolic cognition could only be achieved if sensory and motor information is transformed into a qualitatively different format [Pylyshyn, 1984]. Specifically, this framework proposed that concepts aroused through a transduction process by which sensorimotor inputs (i.e., perceptual features of objects, states, and events) which derived from the environment and the body were transformed into a set of arbitrary representations (i.e., symbols) detached from the perceptual, motor, and emotional systems [Collins and Loftus, 1975, Newell and Simon, 1975, Foltz et al., 1998]. For instance, the concept of a ‘cat’ would be represented in an amodal fashion employing prepositional features such as “it is soft” or “it meows,” rather than the sensory, in this case, tactile and auditory experiences of the softness of the fur or the acoustics of meowing. Similar, the concept of a ‘car’ would be represented by prepositional features such as “it is blue” or “it has an engine,” rather than sensorimotor set of features based on the experience of an individual (Figure 1.2 A). This illustrates how concepts and their meanings were believed to be “autonomous” from the body and “amodal,” not containing any modality-specific information per se [Quillian, 1969, Tyler and Moss, 2001, Pylyshyn, 1984]. Crucially, this view is essentially different from true symbolism where a symbol ‘X’ could refer to ‘a cat,’ ‘sea’ or ‘a Ph.D. thesis.’ According to the symbol processing approach, mental representations of the external signified (e.g., the animal cat) would have identical format and syntactical rules of the external signifier (the linguistic sign, e.g., ‘cat’). Linguistic symbols, one could argue, were thus purely abstract bearing only an arbitrary relation to their contents. Language, being represented in the form of linguistic symbols, bared no relationship to the external referents. The role of the mind was to capture the “disembodied” symbols in a propositional way [Levelt, 1993] and manipulate them or sequence according to the rules of a given grammatical system, thus acting similar to an information processing system.Hence, although there existed no clear hypothesis on the neural correlates of conceptual information during tasks such as comprehension of language or thinking, it was speculated that is could be processed within the heteromodal association cortex, while the modality-specific contents (i.e., tactile or auditory representations) would provide auxiliary aids similar to imagery [Mahon and Caramazza, 2009, McClelland and Rogers, 2003].

Although at that time the symbol-processing approach seemed to provide possible tools to study cognition and psychological processes [Johnson-Laird, 1988], questions regarding the mechanisms of the transduction process, and in particular, how the symbolic representations refer to the objects in the real world provoked doubts and confusion. Indeed, the major criticism or challenge was referred to as the “Symbol Grounding Problem” [Harnad, 1990], which was well illustrated by the so-called Chinese Room Argument [Searle, 1980] (see Figure 1.3). In this thought experiment, an English speaker was closed in a room where s/he received Chinese symbols through a hatch. Their role was to return other Chinese characters following previously explained strict rules; however, without knowing the meaning of the character strings. Critically, if symbols were not causally linked to their referents, the internal manipulation of those symbols would never be sufficient to extract the meaning of the strings. Consequently, the subjects would never understand what information the messages convey, which undoubtedly is not the case of our symbolic or linguistic processing.

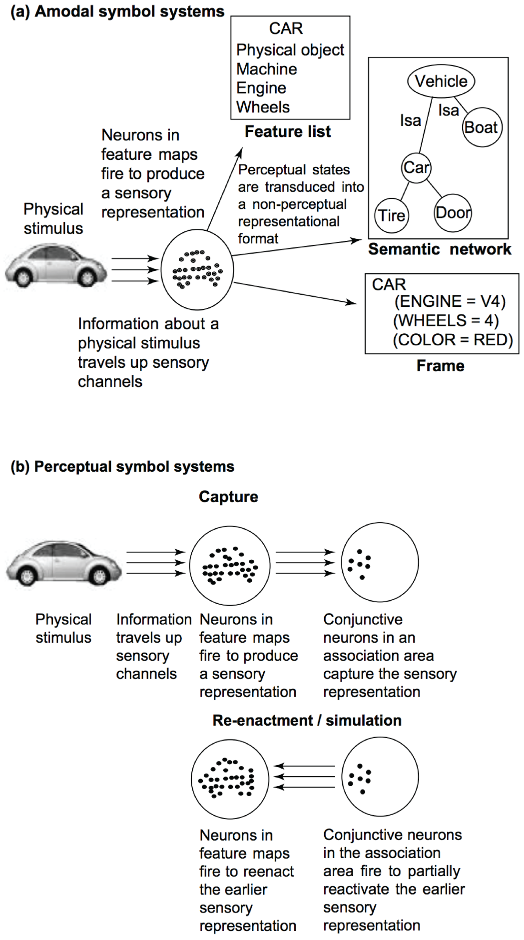


Figure 1.2: (A) Neural representations are established to represent objects in vision. Later, these representations are transduced into a linguistic representation that is amodal (feature list, semantic network or frame). (B) Visual representations are partially captured by conjunctive neurons in nearby association areas. Activating these neurons partially re-enacts the earlier visual states. Analogous accounts exist for acquiring knowledge in the other modalities (adapted from [Barsalou et al., 2003]).

1.2.2 Embodied View of Concepts

In contrast to the theoretical framework of rationalism as well as the traditional views of contemporary cognitive science, is the theoretical framework of Aristotle, and more recent empiricist philosophers such as Locke and Hume, who favored the notion that all relevant knowledge, including concepts and language, emerges from sensory experiences, the so-called impressions, and their copies, namely, the ideas. Their argument asserted that the content of semantic representation *is* the sensory and motor information and that non-perceptual and non-motor reason-based concepts could not be, in principle, informative of either social or physical aspects of the world. The classical view of concepts is also challenged by modern theoretical approaches that embrace the modality-specific view, in line with the thesis of the embodiment, and which supports a tight coupling between the sensory, motor and conceptual brain systems. According to this view, all concepts are fundamentally grounded in the sensorimotor processes emerging from the interactions of an agent with the environment [Warrington and McCarthy, 1987, Lakoff and Johnson, 1999, Martin and Chao, 2001, Barsalou et al., 2003, Pulvermüller, 2005, Gallese and Lakoff, 2005, Kiefer and Pulvermüller, 2012]. Here, all reafferent and exafferent, exteroceptive and interoceptive signals critically constitute meaning and, as such, play an essential role in conceptual thinking and linguistic processing.

In the embodied view, objects in the world are represented in the form of multimodal perceptual symbols which activate associated perceptual and motor information grounded in the history of sensorimotor contingencies of an individual. For instance, the word *cat* is represented as the combination of tactile, visual, auditory, and motor (proprioceptive and kinesthetic) information associated to the animal such as the softness and the color of the fur, the sound of meowing, or the gesture of caressing [Barsalou, 1999, Pulvermüller et al., 2001b]. Similar, the concept *car* is represented in terms of its multimodal features (i.e., the visual representation, the sound and more, Figure 1.2 B). Hence, although the relationship between the arbitrary form/signifier (i.e., the sound of a word or a visual sign) and

© Dan Dennett and Neil Cohn www.visuallanguagelab.com

Figure 1.3: A comic representation of John Searle’s Chinese Room thought experiment (from http://www.visuallanguagelab.com).

the concept/signified (i.e., the meaning) remains arbitrary (i.e., “cat” is “kot” in Polish, “gatto” in Italian, and “gato” in Spanish), the internal representation of the referent *cat* is functionally and neuroanatomically grounded in perceptual and motor representations, rather than on abstract or arbitrary ones such those postulated by the rationalist’s transduction process. Moreover, this grounding is true independently of whether an entity is physically present (e.g., cat), imagined (e.g., Cheshire Cat) or even abstract (e.g., intelligence, magic). Thus suggesting that cognitive processing needs to satisfy the following conditions (1) to represent or be able to refer to referents even if these are not present, and (2) to represent or be able to refer to referents that are not material or cannot be experienced through our senses. To support these functions, not only the online cognition (e.g., actually playing with a cat) has to be coupled to the sensorimotor and emotional systems but also the off-line cognition (e.g., imagining playing with a cat) should recruit the same or similar sensorimotor processes of representation during the process of recall. Such proposals have been widely supported by the scientists and philosophers of embodied cognition who consider brain, body, and environment as interconnected dynamical systems where the brain builds an internal representation of the world by combining a stream of exafferent and reafferent sensorimotor signals [Jackendoff and Jackendoff, 2002, Frith, 2013, Lakoff and Johnson, 2008].

Within such framework, it has been proposed that an individual can achieve stable representations or make inferences via the so-called *simulation process*. This process re-creates an experience through the activation of associated perception-action schemas which recruit the same neural mechanisms and circuitry (within the primary sensory and motor areas) that control overt action [Gallese and Goldman, 1998]. A similar proposal, in this case, form language comprehension, is well illustrated by the Indexical Hypothesis of Glenberg [Glenberg and Robertson, 1999, Glenberg and Robertson, 2000, Glenberg and Kaschak, 2002, Glenberg et al., 2009]. This theory aims at linking language, and in particular semantics, to the preparation and planning of action within a given context. It asserts that an individual perceives the meaning of a situation via a set of possible actions which this situation offers, that is, affordances [Glenberg and Kaschak, 2002]. Through this hypothesis, Glenberg and others propose that words are Indexed to objects in the environment such that the format of concepts directly reflects the form of their referents or, as proposed by Barsalou [Barsalou, 1999, Barsalou et al., 2003], the Perceptual Symbols (PcS) (Figure 1.2 B). Affordances naturally emerge from the objects or

PcS whose processing recruits sensory and motor information in a context-dependent manner, and as such simulation allows comprehension. Within this view, language refers to objects, states, and situations or the affordances of objects, states, and situations [Glenberg and Kaschak, 2002]. For instance, if standing in a room with a freshly brewed coffee and a comfortable chaise longue, one understands and interprets the situation in terms of his/her sensorimotor history related to the experience of coffee or hot beverages and chaise longues or seats in general. Similar, the conceptual and linguistic processing in this context is grounded in the objects *at hand* and their affordances.

The hypothesis that cognition, including higher-level cognitive abilities such as construal of meaning and language processing, is *situated* and *embodied* has been recently gaining consensus in the fields of psychology and cognitive science thanks to vast empirical evidence which will be discussed in the following section.

1.2.3 Current Empirical Evidence

A significant number of behavioral, neurophysiological, and neuroimaging studies with healthy volunteers and brain-damaged patients have been addressing the basis of conceptual and linguistic processing, and in particular the role of sensory and motor representations.

Perceptual simulation in language comprehension

One of the predictions of the embodiment theory is that concepts should depend on the history of sensorimotor interactions of an individual with their environment. In a seminal study, Zwann et al. [Zwaan et al., 2002] demonstrated that comprehension of a sentence, also referred to as semantic analysis, automatically activates specific perceptual representations, in this case, the one of *shape*. To this end, participants were presented with the following sentences: (1) "The ranger saw an eagle in the sky" and (2) "The ranger saw an eagle in the tree." The results yielded significantly faster responses to the picture of an eagle with its wings outstretched after the first sentence as compared to the second one. This further suggests the relevance of the context in building symbolic representations, which seem dynamic and flexible. A similar effect was shown for the mental representation of *orientation* [Stanfield and Zwaan, 2001]. Participants were presented with sentences which indicated a particular orientation of an object, for instance: (1) "He hammered the nail into the wall" (which implies a horizontal orientation) and (2) "He hammered the nail into the floor" (which implies a vertical orientation). The results showed significantly faster responses to a picture of a hummer whose orientation matched the one presented in the sentence.

The effect of contextual representations on language comprehension is not limited to shape, but it extends to other perceptual features such as *color*. To analyze whether implicit perceptual information regarding the color of a given object is processed during comprehension, Connell [Connell, 2007] presented participants with sentences which implicitly implied specific color of the object described, such as "Sarah stopped in the woods to pick up a leaf off the ground." Later, subjects were shown an image of the object from the sentence which either matched or mismatched the implied color (i.e., an orange or a green leaf), and were asked to indicate if the presented object was mentioned in the preceding sentence, or not. Indeed, the response times were manipulated by the color property; however, in this case, they were faster when the color did not match. Consistently with the embodied view, this might suggest that colors are represented differently to other perceptual properties including *shape* and *orientation* which are fundamentally stable independently of the context.

Interestingly, the sensorimotor representations at the basis of language comprehension can be further generalized to novel and unseen situations. For example, in another study on language comprehension [Glenberg and Robertson, 2000], participants were presented with the sentences like the following: "Marissa forgot to bring her pillow on her camping trip". Then they were asked to assess the plausibility of the following sentences: (1) "As a substitute for her pillow, she filled up an old sweater with leaves" and (2) "As a substitute for her pillow, she filled up an old sweater with water". The first sentence was judged more sensible than the second one. This result was interpreted in terms of affordances by which a sweater filled up with leaves acting as a pillow is a novel concept which emerged from the history of sensorimotor contingencies and the goals at hand. On a similar vein, it has been shown that subjects rapidly acquire and understand functions of novel tools [Kaschak and Glenberg, 2000]. Finally, Gerlach et al. [Gerlach et al., 2002] investigated the involvement of the left ventral premotor cortex during the categorizations of different words. Significant activity was found for not only tools but also clothes, fruits, and vegetables as compared to animals and other objects that did not afford motor manipulations.

Concepts activate multimodal information

The simulation theory, which proposes that the brain simulates the properties of objects in a predictive manner verifying their consistency, would predict a cost when switching from one modality to another during a simulation. This cost would be similar to the cost of attention-switching during perception [Spence et al., 2001]. For example, when required to verbally assess the consistency of a given fact involving one modality (i.e., leaves rustle) participants answered significantly quicker if the previously verified fact engaged the same perceptual modality (i.e., blenders make noise: auditory modality) rather than a fact which involved a different modality (i.e., cranberries are tart: gustatory modality) [Pecher et al., 2003]. A number of subsequent studies further supported this hypothesis by demonstrating that subjects simulate the content of the sentence and that the mental representations activate neural patterns within domain-specific sensory areas [Marques, 2006, Vermeulen et al., 2007].

The influence of directional perspective

Another aspect of language comprehension which supports the embodied nature of concepts and words is directional perspective, as studied by Borghi et al. [Borghi et al., 2004]. Using a part verification task, the authors studied if *perspective* implied in a sentence influenced language processing. In the first experiment, subjects had to read sentences describing an object or a place from different perspectives such as from the (1) inside ("You are eating in a restaurant"), (2) outside ("You are waiting outside a restaurant"), or (3) mixed ("You are walking toward and entering a restaurant"). Then, they were presented with a concept-noun (i.e., a table or a sign) and asked to indicate if it belonged to the place, or not. A "sign" would form a part of the outside while a "table" would form a part of the inside of a restaurant. First, the authors reported significantly faster responses to the nouns or objects related to the perspective indicated in the previously read sentence. Second, the responses were significantly faster when the perspective of the object and the location presented one after another matched. Finally, within a given perspective (i.e., restaurant) the responses were faster for the near- as compared to the far-objects. Thus, multiple perspectives indicated by the given sentence modulated access to specific conceptual properties. Importantly, a post hoc analysis confirmed that those results could not be explained by semantic associations, which ruled out possible propositional interpretations. In the second experiment, participants were to indicate if a concept-noun presented after the sentence formed a part of the object included in the sentence, or not, by moving the hand upwards or downwards. Results yielded that moving the arm in a direction compatible with the part location (e.g., responding upward to indicate that a car has a roof or downward to verify that a car has wheels) was faster relative to responding in a direction incompatible with the part location. Crucially, the set of sentences chosen for the experiment did not imply any motor action. Altogether, these results support the hypothesis that language comprehension is based on a simulation process that evokes situated and contextual representations grounded in the proprioceptive and motoric experience of the body.

Can concepts differ from person to person?

Intriguing evidence for conceptual embodiment further comes from studies which show that individual differences can shape the patterns of activity evoked in the motor regions during language comprehension. For instance, Willems et al. [Willems et al., 2010] demonstrated on a sample including both right- and left-handed subjects that lexical decisions involving manual action verbs induce activation in the premotor areas that are contralateral to the dominant hand. This outcome strongly supports the notion that motor component of manual-action verb semantics is body specific, and therefore shaped by actions and sensorimotor experience. Along the same line, it was shown [Beilock et al., 2008, Lyons et al., 2010] that experienced players of ice-hockey present significantly stronger activations in the left premotor cortex for sentences related to hockey actions (e.g., "The hockey player held onto the puck") relative to non-hockey players. The authors found no such differences for sentences related to everyday actions common to both groups (e.g., "the individual closed the book"). Thus, the history of sensorimotor experiences, which is different among individuals, diversely influences linguistic perception.

Motor simulation in language comprehension

The embodiment theory of language supports the hypothesis that concepts and sensorimotor representations are tightly coupled in the brain and, therefore, the understanding of words that refer to the physical properties of an object can influence the way an action is executed and vice-versa. A range of kinematics studies, which allow for the quantification of motor activity during language processing, was used to determine the specificity of the simulation for the representation of *size*. In the study of Gentilucci [Gentilucci, 2003], subjects were required to pronounce syllables "ga" or "ba" while observing gestures of a hand grasping objects of two different sizes. Interestingly, the observation of the hand grasp directed to the bigger object resulted in greater lip aperture as well as voice peak amplitude during the utterances. Glover and Dixon [Glover and Dixon, 2002] instructed participants to reach for and grasp objects which had words "Small" and "Large" printed on their respective surfaces. The authors found an effect of the meaning of words on the grip aperture in the early stages of the motor acts, which faded as the hand approached the target, possibly due to the effect of the visual feedback. With a similar goal, Tucker & Ellis [Tucker and Ellis, 2001] presented subjects with words which denoted relatively larger or smaller objects (e.g., apple and grape). The task was to respond to the stimuli by pressing a block with either a power or a precision grip. Results yielded faster responses when making a grasping movement congruent with the previously presented stimulus suggesting that physical properties of a concept influence motor planning and action execution. Klatzky et al. [Klatzky et al., 1989] showed that priming participant with different hand positions (e.g., clenched hand and a hand with the index finger stretched out), facilitated sensibility judgments of sentences such as "Can you squeeze a tomato?" such that the responses were better in the context of consistent (i.e., clenched hand) as compared to the inconsistent hand position. These studies provide empirical evidence favoring the hypothesis that the brain areas involved in the planning and preparation of hand movement (e.g., grasping) partly share the same neural substrate which is involved in language processing.

Another relevant finding which supports the embodied nature of language is the so-called 'action-sentence compatibility effect' [Glenberg and Kaschak, 2002]. At each trial, participants were presented with one of the following sentences: (1) "Close the drawer" or (2) "Open the drawer." Meanwhile, they were asked to determine whether the sentence is meaningful or not by making a motor movement towards (i.e., pulling movement) or away (i.e., pushing movement) from their body. When presented with the first sentence, the time to respond by making a movement towards the body was significantly slower as compared to the time which a pushing movement would take. This result strongly supports the interaction between the semantics of a sentence and the type of motor command. A similar effect was found by Zaan et al. [Zwaan et al., 2004]. Subjects were to listen to sentences which implied movement towards or away from their body, for instance: "The shortstop hurled the softball to you." Then, they were shown two pictures presented one after another: a small ball and then a big ball, or a big ball and then a small ball. The presentations sequence was suggestive of the movement direction. Here, subjects were to indicate whether the two pictures represent the same objects or not by pressing two different keys. The matching condition yielded faster responses. Additionally, hearing motion-related verbs was shown to interfere with visual motion processing [Meteyard et al., 2007] and vice versa [Meteyard et al., 2008]. None of these results could be accounted for by the amodal theory of cognition [Pylyshyn, 1984].

Neurophysiology and imaging studies

So far, we have reviewed a great body of evidence coming from behavioral studies that supports the embodied nature of language. However, further support to this hypothesis comes from neuroscience and the study of the brain network involved in language comprehension and production. Indeed, multiple studies using functional neuroimaging, EEG, TMS (transcranial magnetic stimulation) as well as neurophysiological recordings in humans and non-human primates show that motor processes are automatically engaged when participants perform perceptual and conceptual tasks. These reports constitute significant evidence in support for the embodied cognition hypothesis [Barsalou et al., 2003, Rizzolatti and Craighero, 2004, Gallese and Lakoff, 2005, Pulvermüller, 2005, Boulenger et al., 2006, Martin, 2007]. For example, it has been shown that reading smell-related words (e.g., cinnamon, garlic, jasmine) elicits increased activation in the primary olfactory cortex relative to neutral control stimuli [González et al., 2006]. Pulvermuller et al. [Pulvermüller et al., 2001a] found topographic differences in cerebral activation patterns as measured by the Event-Related Potentials (ERPs) elicited during the reading of verbs denoting arm (i.e., to grasp), leg (i.e., to kick), and lip/tongue (i.e., to lick) actions. In another study [Pulvermüller, 2005], subjects were asked to passively read action words as well as pseudo-words while high temporal resolution brain activity was recorded using MEG (magnetoencephalography). The activation of motor regions during action word processing appeared to be very rapid ranging between 150-200ms after the stimulus onset. Similar results were obtained by a number of subsequent studies [Hauk et al., 2004, Borreggine and Kaschak, 2006, Boulenger et al., 2006, Hoenig et al., 2008, Revill et al., 2008] supporting the interpretation that the reported activations are part of early semantic access rather than a result of post-comprehension processes of, for instance, imagery (see

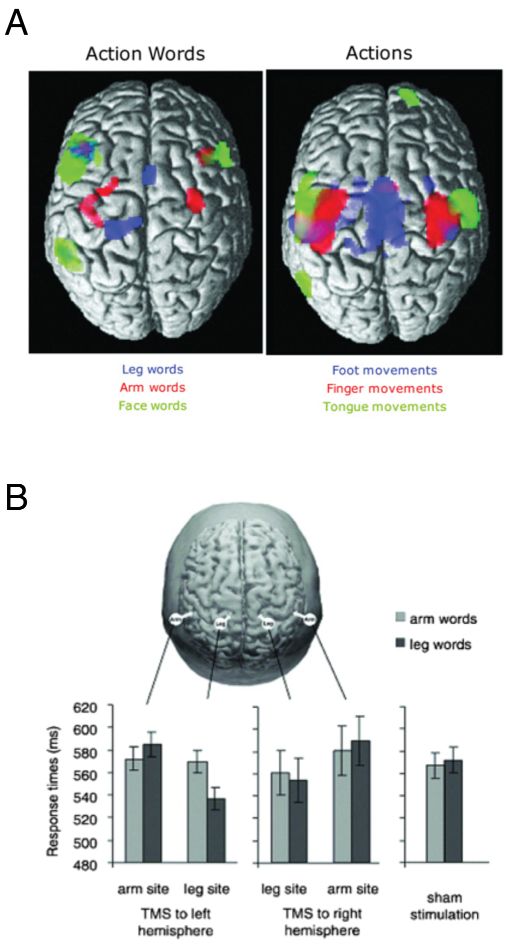


Figure 1.4: (A) fMRI evidence which illustrates overlapping activation during passive reading of face-, arm- and leg-related action words and the actual movements of the same body parts (adapted from [Hauk et al., 2004]). (B) TMS stimulation of premotor arm and leg representations speeds up the recognition of arm- and leg-related words suggesting a causal influence of motor activity on the processing of action words (adapted from [Pulvermüller et al., 2005]).

Figure 1.4 A). Additionally, recent evidence demonstrates that reaching movements are also activated by exposure to action verbs at subliminal levels, that is without them being consciously perceived [Boulenger et al., 2008b], further supporting that semantic representations require modality-specific content.

Finally, TMS studies, which induce transient inactivation in the primary motor cortex or the inferior parietal lobe, complement these results by showing that for instance motor evoked potentials (MEPs) are modulated in an effector-specific manner [Buccino et al., 2005], or that motor activity has a causal influence the processing of action words (Figure 1.4 B). In sum, all these evidence seem to converge to support the coupling of motor and conceptual or perceptual representations in the brain [Oliveri et al., 2004, Pulvermüller et al., 2005, Buccino et al., 2005, Pobric et al., 2010, Ishibashi et al., 2011].

Abstract concepts

A significant challenge for the embodied cognition of language is to explain the origin of abstract concepts, namely, those concepts which are not directly related to sensorimotor representations. How therefore abstract concepts such as *love* can be accounted for within the embodiment framework?

The embodied accounts of higher-level cognitive functions view everyday abstract concepts as fundamentally grounded in action, perception, and emotion such that their referents can be experienced within concrete situations [Barsalou and Wiemer-Hastings, 2005]. For example, the word "satisfaction" might be referred to in a situation, when a fan of Talking Heads band hears the *Psycho Killer* song on a concert. In such context, the concept "satisfaction" is defined by the sensorimotor as well as emotional properties of the experience of listening to the favorite song. Interestingly, there has been evidence in favor of this proposal reported in neuroimaging studies [Pexman et al., 2007, Wilson-Mendenhall et al., 2013]. Specifically, results show significant activity in sensorimotor and emotional brain areas during the processing of abstract concepts.

Another influential hypothesis proposed by Lakoff [Lakoff and Johnson, 2008] is that of conceptual metaphors. Here, abstract concepts or domains such as, for instance, *love* are grounded in the sensorimotor systems through the mediation of concrete domains such as *journey*. Thus, if we consider a metaphor *Love is Journey*, its mapping implies that lovers will correspond to travelers, relationship to the vehicle, and the common goals to the destinations of the journey. Another example could be the one of *Justice is Balance*. We construct an understanding of balance through the sensorimotor interactions with the world, for instance, when walking, balancing a tray or a sea-saw (see Figure 1.5). The resulting schema might be later applied to the understanding of more abstract domains such as justice. For example, we say: "The punishment balanced the crime", "After he paid the fine, they were even," "The fine corresponded to the damage." Therefore, giving a new name to an entity implies providing it with the entire set of analogies, and it is the knowledge of the concrete source domain which is built through the experience of an individual that makes it plausible to reason about the abstract target domain. One of the critical argument supporting this hypothesis is that, indeed, such metaphors are pervasive in all languages.

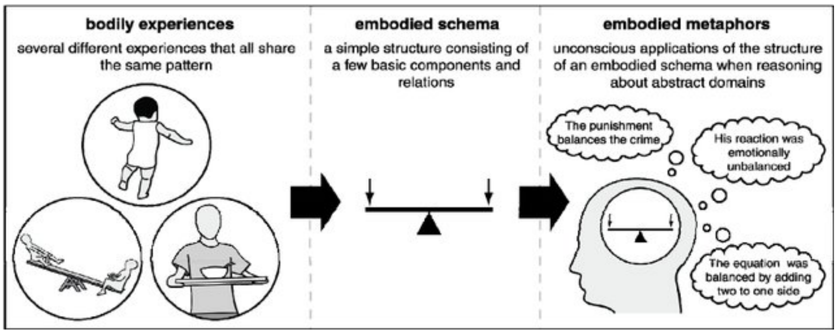


Figure 1.5: Embodied schema and metaphorically related concepts of balance (from [Antle et al., 2009]).

1.2.4 Can We Use the Embodiment Principles to Repair the Brain?

As discussed in the previous section, compelling evidence supports the notion that sensory and motor systems play a critical role in higher-level cognitive functions such as the construal of meaning and language processing [Barsalou, 2008, Lakoff and Johnson, 2008, Pulvermüller, 2005, Glenberg and Kaschak, 2002]. These functions are even more evident when we look at clinical cases where language was lost following specific neurological diseases.

For example, we find evidence from clinical studies which indicate selective impairments in language comprehension in neurological patients suffering progressive supranuclear palsy [Bak et al., 2005], motor neuron disease [Grossman et al., 2008, Bak and Hodges, 2004], apraxia [Buxbaum and Saffran, 2002], as well as Parkinson’s disease [Boulenger et al., 2008a]. In the context of language disorders per se, Arevalo et al. [Arévalo et al., 2007] demonstrated that individuals with nonfluent aphasia showed a selective deficit for both verbs and nouns which related to concepts involving an action with a hand.

Altogether, these findings not only contribute to basic research and the understanding of the origins of linguistic processes but also critically challenge traditional approaches to, among others, first language learning and rehabilitation. For the purpose of this dissertation, we will focus on the possible implications of the aforementioned embodiment hypotheses in the context of remedial work with adult patients whose language had been impaired as a consequence of stroke.

Post-stroke aphasia and traditional language rehabilitation methods

Stroke is a neurological disease which causes the most common disabling functional damages [Carter et al., 2012]. Aphasia occurs in one-third of stroke survivors [Elman et al., 2000]. It is a speech and language disorder which is usually classified according to the fluency, expression, comprehension, repetition, and naming abilities of an affected individual [García-Albea et al., 1996]. Traditional approaches to the treatment of aphasia include both compensatory and facilitatory methods or their combination. The former focuses on learning alternative (often nonverbal) strategies aiming to aid daily communication. Thus, patients train gesturing, drawing, reducing communicative complexity by uttering keywords, and similar [Rogers et al., 1999]. Facilitatory rehabilitation techniques (i.e., restorative treatment), on the other hand, are impairment-based targeting an individual’s linguistic deficits in an individual therapist-patient setting. For word-finding difficulties are the most pervasive symptoms of language impairment, all standard protocols commonly include naming, spelling, repetition, and articulation.

Interestingly, while the compensatory approach does not really aim at the rehabilitation of linguistic skill but rather learning compensatory strategies to overcome communication flaws, the facilitatory approach promotes language practice in isolation from perceptual and motor systems. Focusing primarily on the repetition of words, standard therapies do not put much emphasis on the relevance of multimodal behavioral context in language practice, thus theoretically following the propositional theory which views concepts as abstracted from the body.

A novel approach to language rehabilitation based on sensory-motor learning

Growing body of empirical evidence discussed in the previous section broadens the view on the principles of language (re)learning and challenges the standard language therapy in multiple ways. First, if the premises of the embodiment thesis are true and learning of the words depends on sensorimotor interactions, this would raise the question of whether embedding aphasia rehabilitation training in a goal-oriented action-context, where language is practiced as a behaviorally relevant mean of communication could positively impact recovery. This could occur through, for instance, increased activity in the associated sensory and motor regions, possibly triggering plasticity within the distributed language network. Ineed, a growing body of evidence supports this possibility and a number of studies have already been done in this direction, and they show positive results [Pulvermüller et al., 2016, Difrancesco et al., 2012, Stahl et al., 2016].

Within the scope of this dissertation, we further tested embodiment-based approach to language (re)learning by implementing a multimodal goal-oriented aphasia therapy. We tested the effectiveness of the proposed virtual-reality based technique in the context of a Randomised Controlled Trial (RTC) with chronic patients suffering from nonfluent aphasia. First, we evaluated the safety and usability of the system, which is presented in the first study of Part II of this thesis. Later, we evaluated effects of this therapy as compared to a standard language therapy with 17 patients in a longitudinal clinical trial. The related work is presented in the second study of Part II of this thesis.

Secondly, if as proposed by embodiment the mental representations of concepts are multimodal, that is, they involve multiple sensory (i.e., shape, smell, tactile sensation or sound) and motor (i.e., movement related to its affordance, articulation of the word) features which define it, then providing any of the associated sensory features might facilitate word retrieval in patients with nonfluent aphasia. To explicitly address this question, during a longitudinal RCT, we provided half of the trained stimuli with sensory (i.e., auditory or visual) cues and measured whether the time to utter the word is faster in the condition where the cue is presented. Preliminary results of this study are presented in the third study of Part II of this thesis while full results are summarized in the fourth study of Part II of this thesis.

1.3 Thesis Outline

1.3.1 Part 1: Body Ownership

The fist part of this dissertation is devoted to the study of low-level *minimal* construal of self, in particular, body ownership from the perspective of embodied cognition.

In the first contribution of this dissertation (Chapter 2) [Grechuta et al., 2017b], we aimed at determining whether the structural overlap between the areas underlying the experience of body ownership and motor control is of any functional significance. Specifically, we investigated the question of whether experimentally induced ownership over a fake virtual limb can modulate the performance on a simple sensorimotor decision-making task. The protocol was implemented as a virtual reality-based adaptation of the classical Rubber Hand Illusion (RHI). Participants were randomly split into two conditions where they experienced visual and tactile strokes on their real hand occluded to vision and a rubber analog rendered in front of them in a plausible position. Depending on the condition, the visual feedback of the stroking was either congruent or incongruent with the tactile cues. Meanwhile, they were asked to provide motor responses as soon as they either saw or felt the brush stroking the real hand. We used reaction times to quantify motor performance. Self-reports, as well as galvanic skin responses, served as proxies to the subjective and objective experience of ownership, respectively. The results of this first study demonstrated that body ownership is coupled to behavior such that, during the proposed task, the degree of ownership is correlated with motor performance. This might suggest that body ownership is not exclusively a perceptual and subjective multimodal state, as supported by the classical interpretations of the RHI, but instead it is tightly coupled to systems for decision-making and motor control. We speculated that the coherence of the body might be generated through *agency* (prediction and evaluation of the “reafference”) which might lead to the reorganization and maintenance of body representations (right insular cortex and the frontoparietal circuitry).

The second study (Chapter 3) [Grechuta et al., 2019d] was inspired by the results of the first experiment as well as the empirical evidence suggesting that, when moving in a goal-oriented manner, body ownership is weighted stronger by the congruency of the internal (forward) model of the action and the action-effects rather than the (generative) model of the body and its physical specifics. In particular, we aimed at examining if body ownership can be compromised when action-driven and task-relevant purely distal cues driven by goal-oriented movements do not match sensory predictions. The secondary goal was to determine whether the same manipulation of action-outcomes affects performance in a goal-oriented task (i.e., Air Hockey). We implemented the protocol in virtual reality and manipulated action-driven, purely distal (visuoauditory) cues pertaining to the Air Hockey task. We measured scores, directional error, and reaction times to quantify performance and galvanic skin response, proprioceptive drift as well as self-reports as proxies to the experience of body ownership. Results from all the ownership measures yielded that purely distal cues which pertain to the task and violate predictions about the auditory action outcomes compromise body ownership. Correlations between the proposed measures supported the consistency of the obtained results within three dimensions of ownership quantification including physiological response, behavioral proprioceptive re-calibration, and a conscious report. Additionally, we found that the directional errors were significantly hampered in the incongruent as compared to the congruent condition, which did not depend on the reaction times. These results suggest that the ownership of a body might be driven by bottom-up integration and top-down prediction of purely distal modalities occurring outside of the body and outside of the peripersonal space. We thus speculated that the manipulation of the proposed signals might have reflected on the errors of the forward models which influenced performance and possibly body ownership. This outcome could further suggest that body ownership requires neither the physical or spacial plausibility of the effector, but rather it depends on the sensory prediction errors or, in other words, control, which was the hypothesis of the subsequent experiment.

The main goal of the third contribution (Chapter 4) [De la Torre et al., 2019] was to empirically determine what are the physical and spatial constraints of body ownership in the context of goal-oriented action. First, it was shown that visuomotor synchronicity contributes significantly more to the experience of ownership than visuotactile synchronicity. Moreover, when moving in a goal-oriented manner body ownership seems to be weighted stronger by the congruency of the internal (forward) model than the prior knowledge about the properties of the body. This might suggest that any controllable object (effector) or event could become part of one’s body, provided that its movements are synchronized with self-generated reafferent information intended by the individual. In this study, we explicitly addressed this question. Specifically, we investigated whether controllability over a non-corporeal object located outside of the peripersonal space is enough to induce ownership over this object. We measured galvanic skin response to an unexpected event when the effector was threatened as well as self-reports to objectively and subjectively quantify the experience of ownership. The results yielded significant differences in both self-reports and galvanic skin responses between the two conditions where the effector (a log) was controllable as compared to the condition in which the behavior of the log could not be predicted. The results of this third study are in favor of the proposal that body ownership is a matter of control, not connectedness, space or the physical properties of the body, which sheds new light on the constraints of body ownership in the context of active control.

The final, fourth, contribution (Chapter 5) [Grechuta et al., 2019c] related to the study of the minimal self, and in particular, the feeling of ownership, aimed to determine whether body ownership could, as any coherent percept, depend on the bottom-up integration and top-down prediction of *all* integrated sensory stimuli. In particular, we asked whether ownership can be compromised by not only proximal and distal action-driven cues which are relevant to the task, as shown in the previous studies but also those which pertain purely to the environment. To this end, we created an embodied goal-oriented task (Air Hockey) and manipulated the predictability of the surrounding environment by randomly changing its rules while preserving bodily and action-driven signals entirely predictable. We measured scores, directional error, reaction times, and response times as the proxy to performance. Galvanic skin response, self-reports, and hand withdrawal served as objective, subjective and behavioral measures of ownership, respectively. The results yielded that the way in which we represent our body is indeed contingent upon all the sensory stimuli including purely distal and action-independent signals which belong to the external environment. Thus consistent with the hypothesis of embodiment, it seems that similar to any robust percept, body ownership depends on the consistency of the internal models of not only the body or the consequences of its actions but also the model of the surrounding environment and the rules which underlie its behavior.

1.3.2 Part 2: Language (re)Learning

The goal of the second part of this dissertation was to understand the principles of higher-level cognitive functions (i.e., the construal of the meaning), and in particular language (re)learning, within the framework of embodied cognition and recent neuroscience.

As the first step (Chapter 6) [Grechuta et al., 2016a], we designed and implemented a virtual reality-based language rehabilitation paradigm grounded in the principles of intensity, use-dependent learning and behavioral relevance inspired by the Intensive Language-Action Therapy. Within the context of this first pilot study, we aimed to evaluate the usability and safety of the proposed therapeutic paradigm with two post-stroke chronic patients who suffered from nonfluent aphasia. Besides, before and after a five-day intervention, we evaluated potential changes in linguistic skills of the participants using standard scales that is the Communicative Activity Log (CAL) as well as excerpts from the Western Aphasia Battery (WAB). Crucially, the patients learned how to interact within the virtual-reality interface, and the results yielded high approval and acceptance of the proposed system. Moreover, the scores obtained from the CAL evaluation showed that both the patients and their therapist perceived improvements in the communicative skills after the end of the intervention. Positive changes on the WAB scale further supported this result. We could not exclude, however, that the outcome was biased by the novelty of the system and the peer-peer design of the therapy. Thus to corroborate the findings of this study we designed a longitudinal Randomized Clinical Trial (RCT) with seventeen patients which is the second contribution within the scope of the study of language rehabilitation.

The goal of this contribution (Chapter 7) [Grechuta et al., 2019a] was to systematically investigate whether a virtual reality-based treatment, inspired by the thesis of embodiment and grounded in the principles of use-dependent learning, behavioral relevance, and intensity, positively impacts the recovery from nonfluent aphasia, as compared to a standard speech and language therapy. To this aim, we designed and performed a longitudinal RTC with seventeen post-stroke patients with nonfluent aphasia. We assigned the participants to either the control group which received the standard treatment, provided by the public health-care, or to the experimental group which received augmented embodied therapy with the proposed method called the Rehabilitation Gaming System for aphasia (RGSa). The therapeutic sessions of the RGSa were monitored by a therapy assistant who supervised the patients without offering any elements of the standard therapy. We ensured that both interventions were matched for intensity and materials. Our results demonstrated the effectiveness of RGSa for improving language and communication in all patients who underwent the treatment. Interestingly, the follow-up assessment showed that while both groups retained the vocabulary-related changed only the RGSa group showed therapy-induced improvements in language and communication two months after the therapy. This might suggest that challenges faced by the health-care system in the treatment of stroke can be effectively addressed by augmenting traditional rehabilitation with principle-based computer methods.

The scope of the third (Chapter 8) [Grechuta et al., 2017a] and fourth (Chapter 9) [Grechuta et al., 2019b] contributions related to the principles of language (re)learning was the study of multisensory cueing for lexical retrieval in chronic patients with nonfluent aphasia. Naming disorders happen to be the most frequent symptoms in all types of aphasia and, therefore, the need for word-finding methods is pervasive both within and outside of the clinical context. Traditional cueing methods were shown to have beneficial effects in adding lexical access and verbal execution, however, they often require supervision from a specialist. Grounded in recent neurophysiological findings which support a multimodal language network at the basis of word-meaning, the goal of the last two studies presented in this thesis, was to propose and validate novel multisensory cueing methods to aid word-finding difficulties using a virtual-reality based rehabilitation paradigm (RGSa). Our hypothesis was based on the embodiment thesis and recent scientific findings which support the notion that concepts are multimodal and thus their processing includes visual auditory, gustatory, and olfactory information. Thus, the cues were integrated within a longitudinal clinical trial described in the previous study [Grechuta et al., 2019a] and presented in the form of semantically related sounds (telephone for the phone) or silent videos showing lips articulating the target word. Patients received the cues in half of the practiced stimuli, five days a week, for two months. We used Interaction Times stored in RGSa as a proxy to verbal improvement. In the third contribution, we describe partial results [Grechuta et al., 2017a] about the visual cues. The fourth study comprises full results including both the visual and the auditory cues. Altogether, the results provided clinical evidence for beneficial effects of the proposed multisensory cueing on word naming, especially at the early intervention days when the exposure to the target lexicon was still infrequent. Such findings are in line with the empirical bases of the embodiment hypothesis and suggest that integrating the proposed cues in the rehabilitation of aphasia might foster language-production skills even at the chronic stages of the disease.

Part I

Chapter 2

VISUOTACTILE INTEGRATION MODULATES MOTOR PERFORMANCE IN A PERCEPTUAL DECISION-MAKING TASK

This chapter is based on:
Grechuta, K., Guga, J., Maffei, G., Ballester, B. R., & Verschure, P. F. (2017). Visuotactile integration modulates motor performance in a perceptual decision-making task. *Scientific reports, 7*(1), 3333.

Body ownership is critically dependent on multimodal integration as for instance revealed in the Rubber Hand Illusion (RHI) and a number of studies which have addressed the neural correlates of the processes underlying this phenomenon. Both experimental and clinical research have shown that the structures underlying body ownership seem to significantly overlap with those of motor control including the parietal and ventral premotor cortices, Temporal Parietal Junction (TPJ) and the insula. This raises the question of whether this structural overlap between body owner-

ship and motor control structures is of any functional significance. Here, we investigate the specific question of whether experimentally induced ownership over a virtual limb can modulate the performance of that limb in a simple sensorimotor task. Using a Virtual reality (VR) environment we modulate body ownership in three experimental conditions with respect to the (in)congruence of stimulus configurations. Our results show that the degree of ownership directly modulates motor performance. This implies that body ownership is not exclusively a perceptual and/or subjective multimodal state but that it is tightly coupled to systems for decision-making and motor control.

2.1 Introduction

In order to successfully act in the world, the brain needs to not only process relevant information about the environment but also store and continuously update the position, rotation and velocity of different parts of the body [Ernst and Bülthoff, 2004, Makin et al., 2008]. A simple task, such as intercepting a ball, in practice, requires a number of parallel processes pertaining to both the body and the external world, e.g. postural changes or reaching manipulation [Graziano and Botvinick, 2002, Sober and Sabes, 2003]. Although it has been suggested that much of kinematic control happens outside of perceptual awareness [Johnson and Haggard, 2005], we can be aware of motion, as opposed to immobility, of different parts of the body even when performing automatic movements [Berti et al., 2005, Castiello et al., 1991, Johnson et al., 2002]. This is by virtue of the internal representation of the body [Maravita et al., 2003, Stein and Stanford, 2008], conventionally referred to as body ownership [Head and Holmes, 1911, Tsakiris et al., 2007]. Body ownership accounts for the sensory experiences unique to oneself [Gallagher, 2000, Tsakiris et al., 2010] and it results from the integration of somatosensory and vestibular inputs [Blanke, 2012]. Similarly to the internal models underlying motor control [Ito, 2008, Kawato, 1999, Wolpert et al., 1995], body ownership is subject to multimodal integration [Körding and Wolpert, 2006, Tsakiris et al., 2008] and can be experimentally manipulated [Blanke, 2012, Ehrsson et al., 2004, Tsakiris and Haggard, 2005]. Interestingly, both experimental and clinical research demonstrate that neural substrates for body ownership and internal models driving fine motor control seem anatomically coupled [Berti et al., 2005, Ehrsson et al., 2004, Tsakiris et al., 2007].

Temporal plasticity of the body ownership with regards to the respective roles of vision, proprioception, and touch has been studied experimentally in healthy subjects [Botvinick and Cohen, 1998] using the so-called Rubber Hand Illusion (RHI) paradigm. In this experimental setup, the participants were to view a fake hand being stroked in congruence with tactile inputs provided to their real hand, which was visually occluded. The results suggest that the perception of congruent visuotactile stimulation temporarily modulates body ownership resulting in the experience of ownership of the fake hand. This does not occur when conflicting, incongruent visuotactile inputs are provided. Functional Magnetic Resonance (fMRI) and Positron Emission Tomography (PET) studies investigating the neural correlates of sensory integration driving body ownership, demonstrate that RHI correlates with activity in bilateral premotor cortex (PMC), intraparietal sulcus (IPS), sensorimotor cortex, temporo-parietal junction (TPJ) and the right posterior insula [Ehrsson et al., 2005, Tsakiris et al., 2007, Ehrsson et al., 2004]. Indeed, right insular activity had already been reported in the processing [Vogeley et al., 2004], attribution [Farrer and Frith, 2002] and recognition of the self [Devue et al., 2007] as well as the experience of agency [Farrer et al., 2003]. In a later study, Gentile and colleagues [Gentile et al., 2011] further validated the multisensory integration hypothesis for bodily self-attribution by comparing the properties of regions which are active during visuotactile unisensory and multisensory stimulations of both real and fake hands using fMRI. In both conditions, the authors found activity in the premotor cortices, the insula and subcortical regions, including the right cerebellum and the left thalamus. Coherent with previous literature, these results suggest that underlying the experience of ownership is a set of regions involved in the recognition of self, such as the insula and TPJ; and motor planning, premotor cortices. Tsakiris and colleagues [Tsakiris et al., 2008] showed that the right TPJ correlates with the ability to distinguish self-related events from those generated by the outside world suggesting that it establishes a frame of reference for ownership. Taken together, a number of experimental studies support the notion that body ownership is derived from multisensory integration [Botvinick and Cohen, 1998, Ehrsson et al., 2004, Tsakiris and Haggard, 2005] which correlates with activity in brain structures pertaining to both sensory processing and motor control.

Results from clinical studies investigating pathologies characterized by disturbances of body ownership have provided further evidence for the overlap between sensory and motor areas in body ownership. Somatoparaphrenia, the denial of ownership of a limb or an entire body side, is a consequence of lesions in the right Temporo-Parietal Junction (rTPJ), insula and subcortical regions including the basal ganglia and cerebellum [Vallar and Ronchi, 2009]. Furthermore, anosognosia, or the denial of a diagnosed post-stroke motor or sensory impairment, often follows damage to the insula (hyperacute stages), premotor cortex, cingulate gyrus, and TPJ (subacute stages) [Craig and Craig, 2009, Berti et al., 2005, Karnath et al., 2005]. Interestingly, in case of these acquired neurological pathologies, body ownership disorders are often accompanied by contralesional hemiparesis which might disturb all elements of motor control including decision making, planning and action execution [Baier and Karnath, 2008, Berti et al., 2005, Bourbonnais and Noven, 1989, Vallar and Ronchi, 2009], among others [Coslett, 1998, Haggard, 2005, Pia et al., 2004, Schwoebel et al., 2001, Halligan et al., 1995]. Thus the clinical literature supports that disorders of internal representations of the body might be associated with deficits in motor control, which could result from the overlap of the brain structures involved in the processing of ownership and motor control, in particular the bilateral premotor cortices, TPJ as well as the right insula. Following this line of research, the use of efference copies or chorollary discharge (CP) for sensory input filtering has been proposed as a crucial mechanism for the emergence of the subjective experience of motor control and ownership [Crapse and Sommer, 2008]. Indeed, previous research suggests that clinical conditions leading to ownership delusions, such as neuropathic pain and phantom limb, may relate to defective corollary discharge mechanisms [Ramachandran et al., 1998]. This theory proposes that the accurate virtualization and evaluation of the sensory consequences of self-executed movement may produce the subjective experience of motor control for a specific effector. Little attention has been given, however, to the functional role of this sensorimotor overlap and to the question of whether inducing the experience of ownership may result in a modulation of motor performance.

Here, our goal is to study the relationship between body ownership, decision-making and motor control. In particular, we investigate whether motor performance, in a sensorimotor task, can be modulated by the subjective feeling of ownership over a virtual limb. This modulation is achieved through systematic alteration of the ownership of a virtual arm using the RHI paradigm in Virtual Reality (VR). We devise a protocol to experimentally induce ownership of a virtual hand in healthy subjects, and determine the response times (RTs) in a sensorimotor task where the participants are to deliver rapid motor responses to sensory stimuli (visual or haptic cues) by pressing a button. The degrees of ownership are manipulated across three experimental conditions: congruent visuotactile stimulation (C), incongruent haptic (IH) and incongruent visual (IV) stimulation. Following the RHI paradigm, in the congruent condition, the visuotactile inputs are presented simultaneously, while in the incongruent conditions inputs are delivered asynchronously resulting in visuotactile mismatch. Participants are to respond to the visual and tactile cues in the incongruent visual and the incongruent haptic conditions, respectively. With this design, we on one hand validate previous studies, which found that cross-modal interactions, e.g. haptics and vision, have an effect on the degree of induced ownership, using a VR method [IJsselsteijn et al., 2006, Slater et al., 2008]. Thus we expect that, in the congruent condition, touch is perceived in the location of the virtual hand and the physiological response to an unexpected threatening event presented to the virtual hand is more intense than in IV and IH conditions. Here, we rely on both self-reports [Botvinick and Cohen, 1998] and the Galvanic Skin Responses (GSR) towards a threatening event [Armel and Ramachandran, 2003, Ehrsson et al., 2007]. On the other hand, and most importantly, we analyze whether experimentally induced body ownership driven by visual capture of proprioceptive information modulates motor performance as measured in response times. We expect that in the C motor performance will be faster than both in IV and in IH where the scores will be the lowest. If so, this would suggest a temporal alteration of the internal model that controls overt action possibly deriving form the structural overlap of the brain areas governing sensorimotor processes. Additionally, by deploying two incongruent conditions, where the participants are to rely on either a tactile or a visual cue to execute motor response, we test whether differences in processing of the two sensory stimuli influence the performance on the motor task and whether the sensory weight [Gentile et al., 2011, van Ee et al., 2009] affects physiological responses towards the threat. Here, we expect that in the IV the motor responses may be faster than in IH condition possibly due to the perceptual prominence of vision over touch [Rock and Victor, 1964, Hecht and Reiner, 2009].

2.2 Methods

2.2.1 Participants

Thirty six healthy subjects, from the University campus, were recruited for the study, twenty males (mean age 27.85 ± 4.98) and sixteen females (mean age 26.06 ± 9.55). All the participants were right-handed and reported normal or corrected-to-normal vision. Each of the participants was naÃ¯ve about the purpose of the experiment. Different subjects were randomly assigned to three experimental groups, following a between-subjects design. Such as in [Armel and Ramachandran, 2003], a between-subjects paradigm was chosen to prevent the participants from expecting the threat which could bias the GSR in the subsequent blocks.

The reported experimental procedures with healthy human subjects followed written consents and were in accordance with the established ethical standards, guidelines and regulations. Finally, all the experimental protocols were approved by the University of Pompeu Fabra (Barcelona, Spain).

2.2.2 Experimental Setup

During the experiment, participants were seated at a table with their right palm placed over a fixed point on the table and the left hand placed in a comfortable position at the left side of the table. The ownership was induced to the right hand while the GSR signal and motor responses were delivered by the left hand. Two Ag-AgCl electrodes were attached to the middle and index fingers of the left hand to record the GSR, and the left thumb was placed on the spacebar to deliver motor responses, which prevented from movement artefacts in the GSR trace.

The right virtual hand was displayed in front of the participants in a physically credible position [Salomon et al., 2016], congruent with respect to the real hand (Figure 2.1), through a head-mounted display (HMD, Oculus Rift VR DK1, www.oculus.com). Due to contradictory results from previous studies [Armel and Ramachandran, 2003, Tsakiris and Haggard, 2005, Tsakiris et al., 2010], which indicated a conflict regarding the physical properties (i.e. size, type or weight) of the two limbs necessary to induce ownership, in the present setup we adapted an anatomically plausible virtual hand (Figure 2.1). The tactile stimulation was delivered manually by the experimenter who was seated at the other side of the table, in front of the participant. To fully control for the coincident time onset of the visual (computer generated) and tactile (manually delivered) inputs, the experimenter received precise timing instructions through headphones. For the data analysis, the participants' responses were time locked between the sensory inputs and motor commands, both stored by the system. A paintbrush was used to perform the stroking, and the length of the visuotactile stimulus for every finger was approximately 1.4 seconds long. Virtual analogue of the real brush was accordingly visualized through the HMD.

Within the virtual scene (Figure 2.1), everyone viewed the stimulated virtual analogue of the real right hand which was resting on the table, and the according stimuli in the baseline and intervention blocks. Thus the use of VR allowed us to control for what the participants were exposed to throughout the experiment. In particular, when the Virtual Threat was presented it allowed us to control for unrelated visual factors which could influence the GSR signal. Finally, the present method prevented from possible biases in the performance caused by the presence of the experimenter [Taimela, 1991].

2.2.3 Experimental Protocol

The study consisted of two experimental blocks (Figure 2.1): a baseline block and an intervention block. The first block, or the baseline block (2-3 minutes) was identical for every condition. During this block, the participants were required to provide a motor response as soon as a red sphere appeared in the display in front of them. All the spheres had identical properties and they always appeared in the same position (Figure 2.1, BASELINE). The spheres were displayed with random inter stimulus intervals (1 - 3 seconds) and each exposure lasted 1.4 seconds. This block consisted of 30 trials and served to calculate baseline RTs for every participant. Given the between-subjects design, it allowed us to account for the potential inter-subjects’ variability (i.e. psychophysical differences) and to compare the unbiased motor responses between conditions (i.e. C, IV and IH). The averaged baseline RTs for every participant were later subtracted from the intervention block. Both in the baseline and intervention blocks, the motor responses consisted in pressing the spacebar with the left thumb. Each experimental session in all conditions had an approximate duration of 25 minutes.

The three experimental conditions included C: congruent condition, IV: incongruent visual condition, IH: incongruent haptic condition (Figure 2.1), each of which was followed by a threating event. In these intervention blocks, while the visuotactile stimulation was delivered, the participants were asked to provide a motor response as soon as the right index finger was being stroked. This block consisted of 150 trials and the visual and tactile feedbacks were manipulated across three conditions (Figure 2.1, EXPERIMENTAL BLOCK). In C, the act of stroking with the brush seen on the screen was congruent with tactile stimulation of the real hand. The real finger and the virtual analogue of the same finger were brushed congruently and the participants were instructed to respond to the visual

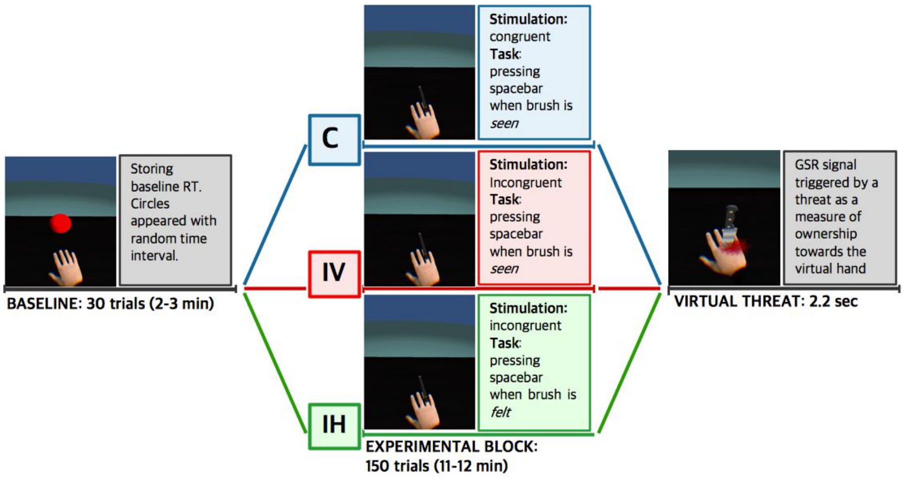


Figure 2.1: The experimental protocol. BASELINE: baseline block (“spheres”). EXPERIMENTAL BLOCK: the intervention block listing each of the three conditions: C: congruent condition (blue), IV: incongruent visual condition (red), IH: incongruent haptic condition (green). VIRTUAL THREAT: measure of physiological responses to a virtual threat. The same colours are used for every condition throughout the article (C: blue, IV: red and IH: green). Baseline block (A) and the virtual threat (C) were the same for every condition.

stimuli (the participants were verbally given the following instruction: “Please, press the spacebar, with your left thumb, as soon as the index finger of the virtual right hand is stroked”). In IV condition, the act of stroking with the brush seen on the screen was incongruent with tactile stimulation of the real hand. The participants viewed a different finger being stroked than the one stroked on the real hand. Here, the participants were instructed to respond only to the visual stimuli (the participants were verbally given the instruction: “Please, press the spacebar, with your left thumb, as soon as you see that the brush strokes the index finger of the right virtual hand”). The IH condition followed the same procedure as the IV, but the participants were asked to respond to haptic stimuli as opposed to visual (“Please, press the spacebar, with your left thumb, as soon as you feel that the brush strokes the index finger of the real right hand"). All the five fingers were brushed in an unpredictable pseudo-randomized sequence, counterbalanced within every session. To account for potential order-effects, we computed a different sequence of strokes for each participant following the same pseudo-random order.

In order to investigate whether the experimentally induced ownership modulates decision-making and motor responses, in every condition, the participants were asked to respond only when the index finger in being stroked. No action was required when the stimulus was provided to other fingers. Furthermore, since the response times can be influenced by instructions emphasizing either speed or accuracy [Zhang and Rowe, 2016], to prevent errors, in our paradigm, the participants were instructed to provide the response when the stroking began, but the task did not impose a speed limit (i.e. no error notification). With such design, at every stroking event, the participants needed to make perceptual decisions of whether to execute the motor action, or not, depending on the visual or tactile inputs provided, while no speed-accuracy trade off was expected. We predicted that in the C condition, with higher ownership, the motor responses will be faster than in both IH and IV. We further hypothesized that the responses in IH might be slower than in IV possibly due to the prominence of vision over touch [Gentile et al., 2011].

At the end of every intervention block, a virtual knife appeared to serve as a threat to the fake hand (Figure 2.1, VIRTUAL THREAT). The knife descended from the top of the screen and into the dorsal part of the virtual hand. The animation gave the impression of the virtual hand being stabbed, which was emphasized by a momentary bout of bleeding emerging from the wound. Both the knife and the blood vanished after less half a second (300 ms.). The whole animation lasted 2.2 seconds in total and the participants were instructed to stay seated, with the HMD on, for another 60 seconds. With this method, we could objectively validate whether synchronous visual and tactile stimulation of the virtual and the real hands can induce the feeling of ownership using the proposed virtual-reality protocol as shown in [Armel and Ramachandran, 2003, Slater, 2009, Rognini et al., 2013]. Secondly, comparison of motor performance and ownership between the conditions, allowed us to investigate our primarily goal, namely, whether the modulation of the representation of the body results in faster responses and a better performance on the proposed motor task. Additionally, by comparing the results from the control conditions (IV and IH) we could further assess whether the performance can be affected by attending to different modalities and whether this influences the GSR responses.

2.2.4 Measures

Self-report

After every experiment, the participants completed a questionnaire, which consisted of nine questions, three of which were related to the perceptual experience of ownership, while the remaining six served as controls. Subjects were asked to respond by rating their level of agreement on a 7-point Likert scale (-3: strongly disagree, 3: strongly agree). The questions were adapted from the previous RHI studies [Botvinick and Cohen, 1998, Ehrsson et al., 2004] to fit the present VR paradigm. The order of the questions was randomized across subjects to avoid order effects. The three questions related to the ownership included: “I had the feeling that I was receiving the touch of the brush in the location of the virtual hand” (Q1), “It seemed as if the touch I felt was caused by the brush that I was seeing on the screen” (Q2), and “I felt as if the virtual hand was my own” (Q3). While the control questions were: “It seemed that my real hand was being displaced towards the left (towards the virtual hand)”, “It seemed that the touch that I was feeling originated in some place in between my own hand and the virtual hand”, “I felt as if my real hand was becoming virtual”, “It seemed (visually) that the virtual hand was being displaced towards the right (towards my real hand)”, “The virtual hand started to look like my own hand in some aspects”, and “I had the sensation of having more than one right hand”.

Galvanic Skin Response (GSR)

The Autonomic Nervous System (ANS) is the primary mechanism which regulates involuntarily physiological states, such as arousal produced due to anticipating pain or fear. We used GSR (the electrical conductance of the skin), as a measure of ANS activity to further quantify the experience of ownership over the virtual hand and compare our results with previous studies [Armel and Ramachandran, 2003]. We expected that all subjects would show changes in GSRs after the threatening event (VT), but that there would be higher responses in the congruent (C) condition due to the enhanced assimilation of the virtual hand into the perceptual bodily representation.

The GSR was recorded throughout the experiment with two Ag-AgCl electrodes attached to the palmar surface of the index and middle fingers of the participants' left hand (e-Health Sensor Platform V2.0, Cooking hacks, Zaragoza, Spain) and the data was recorded using an Arduino microcontroller [Banzi et al., 2013]. We measured GSR during the entire experiment, however, we were particularly interested in the GSR responses to the VT displayed at the end of the experiment in every condition (Figure 2.1, VIRTUAL THREAT). The timing of the threat event was stored and registered with the GSR and behavioral record for further analysis. In order to compare the GSR responses between the three conditions, we defined a latency onset window of 12 seconds after the stimulus onset. The GSR signal after VT was normalized for every participant by subtracting the mean signal from 12 seconds prior to the stimulus onset.

Response Times (RTs)

In the baseline block ("Spheres"), all the participants were asked to provide a motor response (press the spacebar) as soon as a red sphere appeared in front of them. To calculate the baseline and account for individual differences between subjects (i.e. psychophysical inter-subject variability), we stored the RTs for every participant, which we defined as the time interval between the onset of the sphere and motor response. During the intervention block, (Figure 2.1, EXPERIMENTAL BLOCK) the RTs were defined as the intervals between the beginning of stroking and motor response. In both blocks, the RTs were used as a measure of perceptual detection and motor performance [Gentile et al., 2011, Keele, 1986]. For the data analysis, we normalized the RTs in the intervention block for every participant by subtracting their mean response time from the baseline block ("spheres").

2.3 Results

The goal of the present study was first to devise and validate a VR paradigm of the standard RHI protocol [Botvinick and Cohen, 1998, Ehrsson et al., 2004, Makin et al., 2008] following two ownership induction methods (i.e. congruent and incongruent). Second, and most importantly, we evaluated the effect of ownership of the virtual hand, as measured by self-reports [Botvinick and Cohen, 1998] and GSR responses to a virtual threat [Armel and Ramachandran, 2003], on motor performance in the proposed task. We expected that ownership might have a modulatory effect on motor performance as measured through RTs such that in C the performance will be better than in both control conditions (IV and IH). Finally, grounded in the theories of the dominance of vision over touch [Rock and Victor, 1964], we further hypothesized that in IV the motor performance can be faster than in IH condition.

Normality test revealed that GSR and RTs data were not normally distributed. Consequently, the statistical analysis followed nonparametric analysis. We used Kruskal-Wallis (KW) tests between conditions, corrected for multiple comparisons and a Mann-Whitney U test to identify differences between groups. A Pearson product-moment correlation coefficient was computed for the subsequent linear correlation analyses.

2.3.1 Self-reported Feeling of Ownership

After each experimental session, the participants were required to rate the level of perceived ownership. Results show that the congruent visuotactile stimulation of the virtual and real hands (i.e. condition C) enhanced the feeling of ownership, compared to the control conditions IV and IH ($p = 0.016$) (Figure 2.3). The mean score across the participants for Q1, Q2 and Q3 in condition C was 0.78 ($SD = 1.82$), -0.76 ($SD = 1.75$) in IV, and -1.0 ($SD = 1.39$) in IH. Kruskal-Wallis (KW) test showed that there was a significant difference between the three conditions (C, IV, IH) for the three ownership questions ($H(2, 36) = 18.71, p < 0.001$). In particular, we followed the previous finding with a Mann-Whitney U test, which indicated that the scores in condition C were significantly higher than both conditions IV ($Mdn = -1, U = 323, p < 0.001$), and IH ($Mdn = -2, U = 292.5, p < 0.001$). No significant difference was found between the control conditions, IV and IH ($U = 556.5, p = 0.32$) (Figure 2.3). The mean rating across all participants for the six control statements (Q4, Q5, Q6, Q7, Q8, Q9) was -0.36 ($SD = 2.13$) for condition C, -0.75 ($SD = 1.68$) in IV, and -0.61 ($SD = 1.96$) in IH. No difference was found between the conditions in the control questions ($KW, H(2, 36) = 0.81, p = 0.67$) (Figure 2.2). Thus self-reported feeling of ownership occurred only in condition C.

2.3.2 Physiological Measures of Ownership Illusion

We stored and analyzed the GSR (Figure 2.4) as a quantitative measure of ANS to further analyze differences in the induced feeling of ownership across the three experimental conditions. Prior to the data analysis, we calculated the mean GSR, as an integral of the curve in a time window of 12sec, together with its associated standard deviation (SD) for every condition, and excluded three participants whose mean response was 2.5 SDs higher or lower than the mean of the group (2 participants whose signal was higher in C and IH, and 1 participant whose signal was lower in IV) (Figure 2.5). As hypothesized, we found significant differences in the GSR data between the three conditions ($KW, H(2, 32) = 1256.5, p > 0.001$). A Mann-Whitney follow-up test indicated that post-threat event mean GSR responses were significantly higher in C condition ($Mdn = 13.64$) than in IH ($Mdn = 3.51, U = 229265, p < 0.001$) and in IV

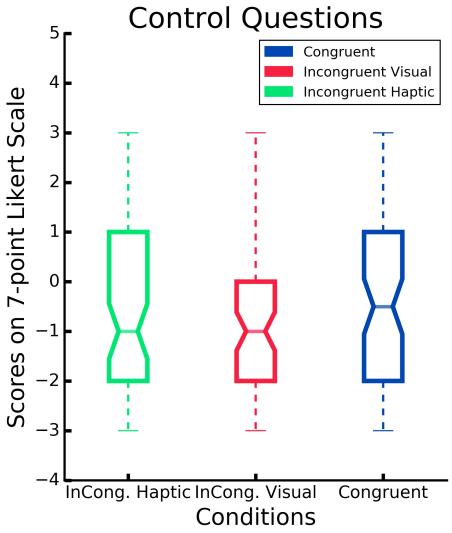


Figure 2.2: Self-reported experience of ownership. Y-axis: Responses on the 7-point Likert scale ranging from -3 (strongly disagree) to 3 (strongly agree). Scores above 0 indicate a feeling of ownership. Control Questions: mean of the three questions related to the ownership illusion per condition.

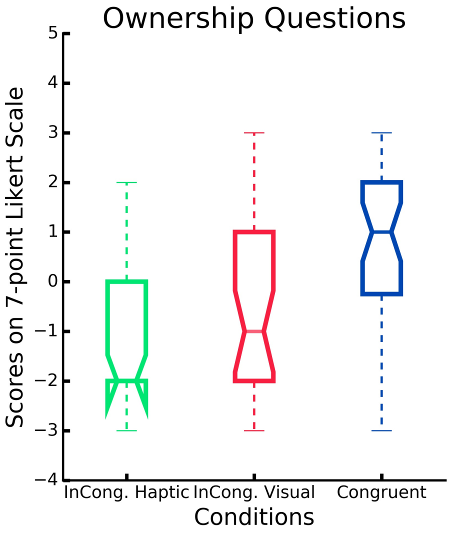


Figure 2.3: Self-reported experience of ownership. Y-axis: Responses on the 7-point Likert scale ranging from -3 (strongly disagree) to 3 (strongly agree). Scores above 0 indicate a feeling of ownership. Ownership Questions: mean of the six control questions per condition.

($Mdn = 7.24, U = 328404, p < 0.001$) conditions. Finally, we observed a significant difference between the control conditions IV and IH ($U = 291783, p < 0.001$). The GSR outcome further validates the results from the self reports suggesting the highest ownership in condition C.

2.3.3 Performance

Reaction times served as the performance measure in the proposed task. Medians per condition prior to normalization are: C ($Mdn = 70.0$),

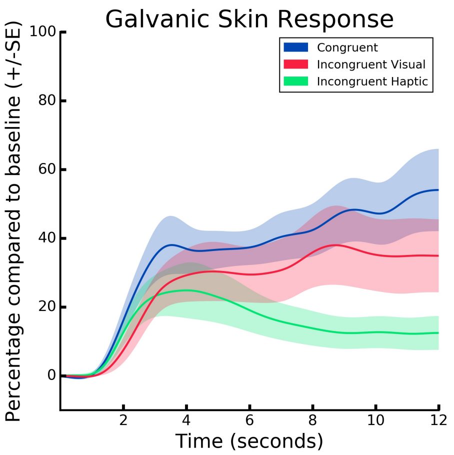


Figure 2.4: GSR results. The sampling rate for the GSR signal was 60Hz. Accordingly, the data were run through a low-pass filter with a cut-off frequency of 0.06Hz. Mean GSR responses per condition for all the participants averaged in a time window of 12 sec. The threatening event happened at time = 0.

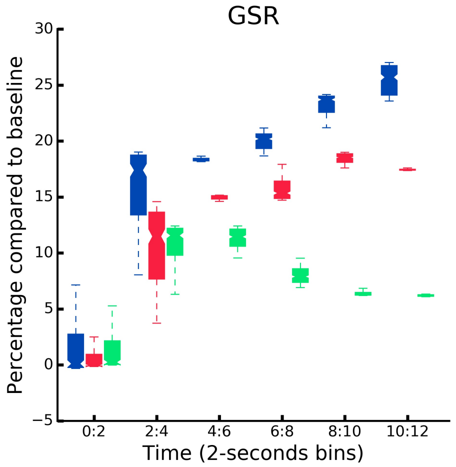


Figure 2.5: GSR responses per condition binned in 2.5-seconds time windows. The first bin (0-2 seconds) represents the latency of the GSR response following the threat ($time = 0$).

IV ($Mdn = 105.0$) and IH ($Mdn = 200$). The RTs measured in the first block ("spheres") served to calculate the baseline (i.e. inter-subjects psychophysical differences) for every participant, which was subtracted from the intervention block for the performance analysis. As expected, in this block, no differences in RTs were found between the three conditions ($KW, H(2, 36) = 3.3, p = 0.19$). In the intervention block, the reaction times served as a measure of motor performance in the proposed task. We observed a significant difference between the three conditions ($KW, H(2, 36) = 896.9, p < 0.001$) (Figure 2.6). A Mann-Whitney test further indicated that the RTs were significantly lower in condition C ($Mdn = 94.16$) than in both condition IV

($Mdn = 129.0$), ($U = 480559.0, p < 0.001$), and in IH ($Mdn = 205.83$), ($U = 216680.5, p < 0.001$). Additionally, we found that the RTs were significantly higher in the IH than in IV ($U = 259008.5, p < 0.001$). In condition IH, we found a significant difference between first, middle and last trials ($KW, H(2, 12) = 6.72, p = 0.034$) (Figure 2.7). No such differences were found in C ($KW, H(2, 12) = 2.34, p = 0.3$) nor IV ($KW, H(2, 12) = 1.04, p = 0.51$) (Figure 2.7). We observe that the congruency of visuotactile stimuli modulated motor responses. We further report a difference in RTs between IV and IH possibly due to sensory predominance of vision over touch.

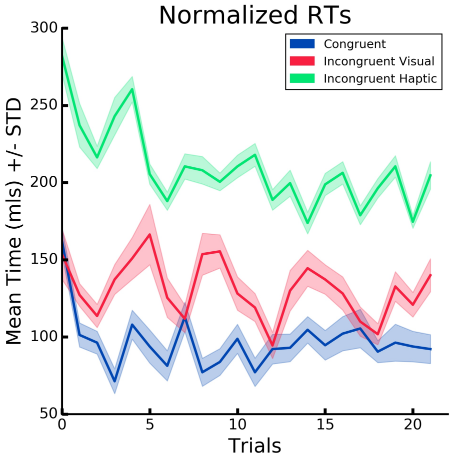


Figure 2.6: Response Times. Normalized mean of the response times for all the participants, defined as the intervals between the start of the stroking and subjects’ response, over time, per condition.

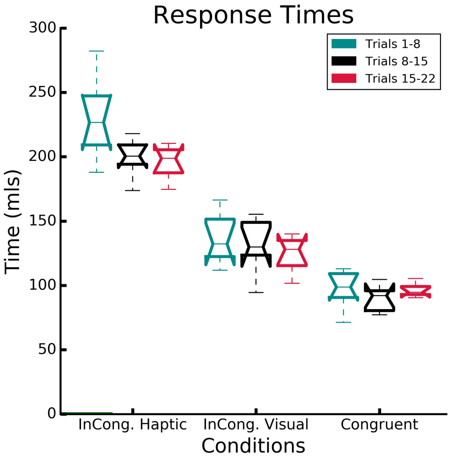


Figure 2.7: Normalized RTs responses for the three conditions binned in windows of 7 trials (1-8: early trials, 8-15: middle trials, 15-22: late trials).

Finally, we computed the error responses in every condition for the experimental block including false positives (i.e. motor response provided given a stimulus to a finger different than the index finger) and anticipatory responses. For all the subjects and all the trials we report 4 errors in C, 3 in IV and 5 in IH. No significant difference was found in the errors’ RTs between the three conditions ($KW, H(2, 36) = 0.85, p = 0.65$). Finally, we report no anticipatory responses in neither of the three conditions.

2.3.4 Correlation Analysis

A Pearson product-moment correlation coefficient was computed to assess the relationship between the ownership measures and performance on the motor task. For the analysis, we computed the mean GSR from the last five seconds post-threat, mean RT of the last 5 trials, and the mean score from the Ownership Questions for each participant in every condition, respectively. We report a significant negative correlations between post-threat GSR responses and the RTs ($r = -.35, p = 0.04$) (Figure 2.8) as well as the ownership questionnaire outcome and the RTs ($r = -.5, p = 0.003$) (Figure 2.10). Finally, the linear relationship was computed between the post-threat GSR responses and the ownership questionnaire. We observed a positive correlation between the two measures ($r = .37, p = 0.03$) (Figure 2.9). This result confirms the consistency of the three different dimensions of ownership measure (i.e. behavioral, conscious report and physiological reaction). In addition, this results might suggest that the feeling of ownership can have different levels on a continuous scale rather than being a binary state.

2.4 Conclusions

The goal of the present study was to investigate whether experimentally induced ownership can result in a modulation of motor performance in the proposed sensorimotor task. We hypothesized that the congruent, spatiotemporal pattern of visuotactile stimulation will account for higher feeling of ownership of a virtual limb and consequently enhanced performance in C as compared to the control conditions. To test this hypothesis we adapted the traditional RHI paradigm in a VR setting. The visual feedback of the active touch was manipulated across three conditions where

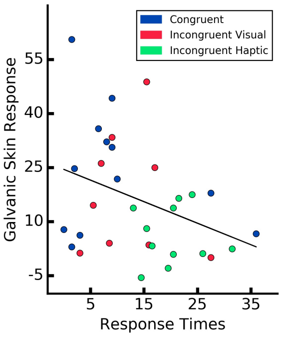


Figure 2.8: Correlation between the performance measure (mean RTs from the last five trials for each participant) and post-threatening galvanic skin response signal (mean GSR signal after the first peak). Every data point represents a participant.

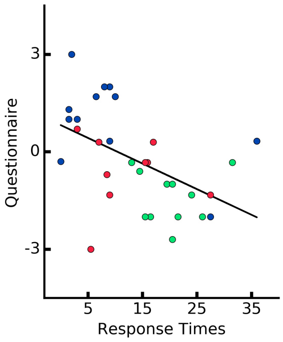


Figure 2.9: Correlation between the motor performance and self-reports (mean of the three ownership questionnaire for each participant).

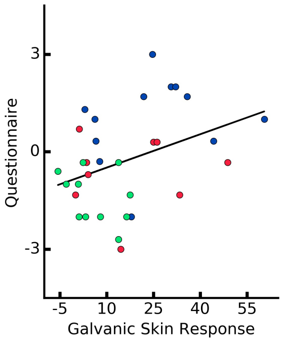


Figure 2.10: Correlation between the post-threatening galvanic skin response signal and self-reports. Every data point represents a participant.

we varied the congruence of inputs and the choice modality.

Results from self-reports suggest that the participants experienced touch in the location of the virtual hand when tactile and visual stimuli were delivered congruently. As expected, the scores in C were significantly different from both IV and IH conditions, where the stimulation was incongruent. To further assess the feeling of ownership and validate the data collected through self-reports, at the end of every experiment, we introduced a virtual threat and computed the GSR responses as an objective measure of autonomic, physiological arousal. We expected that the threatening event would provoke changes in GSR in all the three conditions, but that the subjects experiencing congruent visuotactile feedback would present higher responses due to the enhanced assimilation of the virtual hand into the body representation. Indeed, in C participants showed significantly higher GSR responses than in both IV and IH. Interestingly, we also observe statistically higher GSR responses in the IV than IH condition. Similar significant differences in performance (i.e RTs) were observed between the three conditions such that participants in C responded the fastest, and those in IH the slowest.

To investigate the relation between the degree of ownership, perceptual decision making and motor performance in the proposed task, we performed correlation analyses. We found that both subjective (i.e. self-reports) and objective (i.e. GSR) measures of ownership were significantly correlated which supports the use of present method to measure ownership. Secondly, we report significant correlations between both ownership measures and performance on the motor task. Interestingly, although both in IV and IH conditions the pattern of visuotactile feedback was incongruent and no ownership was expected, in the IV the performance and the GSR were higher than in IH while the measures were correlated. This might suggest that the attended modalities weighted ownership differently, such that despite the congruency of the feedback, when vision was attended to elicit motor action (i.e. IV) the ownership was higher than when the tactile stimuli was attended.

Overall, the reported results are consistent with previous studies within the framework of multimodal processing [Blanke, 2012, Botvinick and Cohen, 1998, Stein and Stanford, 2008] showing that the degree of ownership towards an artificial or, as in our case, virtual hand is associated with crossmodal processes of visuotactile stimulation, which can be measured both subjectively and objectively. Moreover, differences in response latencies between C and IV support that ownership has a modulatory effect on perceptual decision-making processes, which are coupled to behavior (i.e. RTs), physiological processing (i.e. GSR) and conscious perception (self-reports). Finally, our data suggests that ownership might be manipulated not only by the congruence of visuotactile inputs but also by the weight of the sensory stimuli that is being attended.

2.5 Discussion

Accomplishment of even simple behavioral goals requires planning, execution, and complex coordination of movements involving different parts of the body [Ernst and Bülthoff, 2004, Makin et al., 2008]. Thus, in order to successfully interact within the external world, the brain needs to continuously process the information about the body and the surroundings so that it can adjust its internal model [Kawato, 1999] to the environment. The present results, in particular differences in performance between C and IV conditions, can be interpreted in terms of functional role of ownership and are consistent with previous literature in the domain of motor control [Kawato, 1999, Miall and Wolpert, 1996] supporting the idea that body ownership acts as an internal, dynamic model [Stein, 1998] where bodily properties driven by multisensory integration can modulate motor performance. During the experiment, the participants had to make real-time decisions on whether to execute a particular motor action and press the button or not, depending on sensory inputs. When the visuotactile inputs were delivered congruently, such as in C, the amount of sensory information reduced perceptual ambiguity which triggered according motor response faster than when the incongruent stimuli were provided [Samad et al., 2015]. These differences in performance between C, IV and IH seem in line with the Bayesian principles of multimodal integration for decision-making mechanisms [Beck et al., 2008, Körding and Wolpert, 2006, Mamassian et al., 2002, Samad et al., 2015]. Alternatively, as had previous studies found [Gentile et al., 2011], this result may also indicate that congruent combination of information from different sensory modalities (i.e. vision and touch) facilitated the ability to recognize specific sensory stimuli resulting in faster perceptual discrimination and decision-making independent of ownership. Furthermore, during the experimental block subjects could learn to plan their responses differently when facing two asynchronous conditions IV and IH. We exclude, however, that the RTs might have been affected purely by the sensory congruency or expectations, since this effect would not lead to correlations between RTs and measures which did not directly depend on motor planning and motor control (i.e. GSR or self-reports). We believe that the present results highlight the role of ownership in the context of motor actions. We further observe, however, that the visuocactile integration also affected both the autonomous responses, measured through GSR and the conscious perception, measured through self-reports. This might suggest further that multisensory (i.e. visuotactile) integration has multiple dimensions including behavioral [Wolpert et al., 1995, Stein and Stanford, 2008] (i.e. motor), physiological [Armel and Ramachandran, 2003] and conscious [Head and Holmes, 1911, Vogeley et al., 2004, Blanke, 2012, Limanowski, 2014], which is supported by the correlation analysis between the discussed measures.

On the other hand, we observed significant differences in performance between IV and IH. Since the two conditions involved different sensory modalities to elicit motor response (i.e. visual or tactile), the reported result could be possibly explained by differences in terms of tactile and visual processing [Gentile et al., 2011]. Given perceptual dominance of vision over touch, and coherent with literature [Rock and Victor, 1964, Hecht and Reiner, 2009], in the IV the motor responses were faster than in IH condition. Interestingly, however, we further observe a significant difference in GSR responses between IV and IH conditions such that the GSR in the IV was significantly higher than in IH. We elaborated on this result by showing its relationship with the two other measures used in the study: autonomous physiological response and self reports. The reported correlations emerging from this analysis support the hypothesis that significant changes in reaction times across conditions could be effectively due to a modulation of the body representation. Furthermore, the high accuracy in motor performance, equally distributed in every condition, suggests that the difference in reaction times could be due to the same process of bodily representation affecting the subjective feeling of ownership. Since the visuotactile stimuli in IV and IH were incongruent, the feeling of ownership in IV should not have been modulated by sensory integration. Instead, we propose that the feeling of ownership in IV might have been modulated through the weight of the attended modality [Gentile et al., 2011], such that despite the congruency of the feedback, the ownership is higher when visual stimuli are attended to elicit motor action, than when the tactile stimuli are attended.

From the neuroscientific perspective, present results might be explained in terms of the anatomical coupling of brain structures underlying ownership and motor control. Clinical-pathological studies [Berti et al., 2005, Karnath et al., 2005, Vallar and Ronchi, 2009] suggest that disturbances in attitudes towards body ownership tend to overlap with disorders in motor control, pointing to lesions in parietal and ventral premotor cortices, TPJ and the insula. The same set of brain areas have been identified in experimental studies investigating the multisensory nature of body ownership, using fMRI or PET [Berti et al., 2005, Ehrsson et al., 2004, Tsakiris et al., 2007, Vallar and Ronchi, 2009]. Different performance outcomes presented here, seem coherent with this literature suggesting that perceptuomotor abilities result from multisensory integration mechanisms, such as in C condition. These mechanisms generate a coherent reference model and reduce perceptual ambiguity in the moment of decision-making enhancing the motor response. Such motor behavior can be modulated by providing incongruent spatiotemporal sensory cues with varying weights [Gentile et al., 2011], as in IV or IH.

One of the critical aspects of the present study was that the participants were to provide motor responses using the left hand given sensory cues provided to the right hand with the induced ownership. Accordingly, we show that the degree of induced ownership indeed modulated the response times of the motor commands in the left hand. This result supports the theory that the coherence of the body is generated through agency (i.e. an accurate prediction and evaluation of the "reafference") and leads to the reorganization and maintenance of body representations possibly located in the right insular cortex and the frontoparietal circuitry, the neural territories associated with the subjective experience of ownership [Tsakiris et al., 2007]. In their study, Tsakiris and colleagues [Tsakiris et al., 2006b], show that the integration of visuotactile stimulation induces body ownership locally, in a fragmented manner. They further propose, however, a secondary mechanism, possibly a generalization of the visuotactile associations, which accounts for perceiving the body as a coherent entity.

Our results seem in coherence with this hypothesis suggesting that agency might play a role in the proposed secondary mechanism.

Applications of the presented paradigm might have relevance in fields such as motor rehabilitation. Acquired brain lesions including stroke often result in ownership disorders (i.e. anosognosia) and hemiparesis, which impair motor functions of upper extremities [Coslett, 1998, Halligan et al., 1995, Pia et al., 2004, Schwoebel et al., 2001]. Recently, a number of studies examined the functionality of virtual reality based rehabilitation systems that aim at post stroke motor recovery of upper extremities [Jack et al., 2001, Merians et al., 2002, Cameirão et al., 2010, Saposnik et al., 2010]. Some of these setups are designed so that a motion sensor continuously tracks the user’s arms, and the movements are projected into the virtual scenario from a first person’s perspective. The underlying hypothesis for this rehabilitation research is that sensorimotor contingencies build up through experiential learning, and thus follow the statistics of the multimodal inputs that are exposed to the brain triggering plasticity which may lead to recovery [Johansson, 2000, Kleim and Jones, 2008]. Indeed, several studies show promising results [Cameirão et al., 2010, Maier et al., 2014, Ballester et al., 2015c]; however, none of them has explicitly addressed the question of whether inducing ownership towards the virtual effector might be beneficial for the rehabilitation purposes by reinforcing acquired sensorimotor contingencies and subsequently modulating arm use and motor performance. Further clinical studies will be conducted to evaluate whether the present method applies to hemiparetic stroke patients and whether induced ownership using virtual reality may influence recovery processes.

Chapter 3

SELF BEYOND THE BODY: ACTION-DRIVEN AND TASK-RELEVANT PURELY DISTAL CUES MODULATE PERFORMANCE AND BODY OWNERSHIP

This chapter is based on:
Grechuta, K., Ullysse, L., Ballester, B. R., & Verschure, P. F. (2019) Self beyond the body: action-driven and task-relevant purely distal cues modulate performance and body ownership. *Frontiers in Human Neuroscience.*

Our understanding of body ownership largely relies on the so-called Rubber Hand Illusion (RHI). In this paradigm, synchronous stroking of the real and the rubber hands leads to an illusion of ownership of the rubber hand provided that it is physically, anatomically, and spatially plausible. Self-attribution of an artificial hand also occurs during visuomotor synchrony. In particular, participants experience ownership over a virtual or a rubber hand when the visual feedback of self-initiated movements follows the trajectory of the instantiated motor commands, such as in the Virtual Hand Illusion (VHI) or the moving Rubber Hand Illusion (mRHI). Evidence yields that both when the cues are triggered externally (RHI) and when they result from voluntary actions (VHI and mRHI), the experience of ownership is established through bottom-up integration and top-down prediction of proximodistal cues (visuotactile or visuomotor) within the peripersonal space. It seems, however, that depending on whether the sensory signals are externally (RHI) or self-generated (VHI and mRHI), the top-down expectation signals are qualitatively different. On the one hand, in the RHI the sensory correlations are modulated by top-down influences which constitute empirically induced priors related to the internal (generative) model of the body. On the other hand, in the VHI and mRHI body ownership is actively shaped by processes which allow for continuous comparison between the expected and the actual sensory consequences of the actions. Ample research demonstrates that the differential processing of the predicted and the reafferent information is addressed by the central nervous system via an internal (forward) model or corollary discharge. Indeed, results from the VHI and mRHI suggest that, in action-contexts, the mechanism underlying body ownership could be similar to the forward model. Crucially, forward models integrate across all self-generated sensory signals including not only proximodistal (i.e., visuotactile or visuomotor) but also purely distal sensory cues (i.e., visuoauditory). Thus if body ownership results from a consistency of a forward model, it will be affected by the (in)congruency of purely distal cues provided that they inform about action-consequences and are relevant to a goal-oriented task. Specifically, they constitute a corrective error signal. Here, we explicitly addressed this question. To test our hypothesis, we devised an embodied virtual reality-based motor task where action outcomes were signaled by distinct auditory cues. By manipulating the cues with respect to their spatial, temporal and semantic congruency, we show that purely distal (visuoauditory) feedback which violates predictions about action outcomes compromises both performance and body ownership. These results demonstrate, for the first time, that body ownership is influenced by not only externally and self-generated cues which pertain to the body within the peripersonal space but also those arising outside of the body. Hence, during goal-oriented tasks body ownership may result from the consistency of forward models.

3.1 Introduction

Humans and other species simultaneously acquire and integrate both self-generated (i.e., reafferent) and externally-generated (i.e., exafferent) information through different sensory channels [Sperry, 1950]. Hence, the ability of the nervous system to generate unambiguous interpretations about the body, the so-called body ownership, and determine the source and relevance of a given sensation is fundamental in adaptive goal-oriented behavior [Botvinick and Cohen, 1998, Ehrsson, 2012, Wolpert and Flanagan, 2001, Van Den Bos and Jeannerod, 2002]. Imagine playing Air Hockey where the objective is to score points by hitting a puck into the goal. To accomplish the task, at every trial, the brain prepares and generates actions which are most likely to elicit the desired trajectory leading the puck towards the target [Sober and Sabes, 2003, Shadmehr et al., 2010, Wolpert and Flanagan, 2001]. Simultaneously, it predicts the sensory consequences of those actions from proprioceptive or tactile modalities which inform about the position and location of the arm, and from visual or auditory modalities which inform about the position and location of the puck [Miall and Wolpert, 1996, Ernst and Bülthoff, 2004, Makin et al., 2008]. Since both types of cues may constitute a corrective error signal for the consecutive trial, they are both relevant to the task [Shadmehr et al., 2010, Wolpert et al., 2011]. This evidence suggests that the internal models of the external environment, the motor apparatus, and the body are being continuously shaped and updated through sensorimotor interactions of an agent with the world [Miall and Wolpert, 1996, Tsakiris, 2010, Blanke, 2012, Apps and Tsakiris, 2014]. Specifically, this tuning occurs through a combination, integration, and prediction of both reafferent and exafferent signals from multisensory sources [Prinz, 1997, Ernst and Bülthoff, 2004, Noë, 2004]. However, mechanisms driving the representation of self and, in particular, body ownership in action contexts which require manipulation of the environment and therefore integration of not only proximal or proximodistal but also purely distal cues remain elusive.

In fact, our understanding of body ownership largely relies on the so-called Rubber Hand Illusion (RHI) where subjects passively receive sensory stimuli [Botvinick and Cohen, 1998]. RHI is a well-established paradigm [Botvinick and Cohen, 1998, Makin et al., 2008, Tsakiris, 2010] where the illusion of ownership towards a rubber hand emerges during externally-generated synchronous, but not asynchronous, stroking of the real and fake hands [Botvinick and Cohen, 1998]. The illusion generalizes to distinct body-parts including fingers, face or a full body [Dieguez et al., 2009, Sforza et al., 2010, Lenggenhager et al., 2007]. Initially, Botvinick & Cohen [Botvinick and Cohen, 1998] proposed that the illusion of ownership over the rubber hand is a rather passive sensory state which emerges reactively from a bottom-up integration of multisensory, in this case, visuotactile signals (i.e., proximodistal). Interestingly, subsequent studies investigating mechanisms underlying the RHI extended this classical interpretation by demonstrating that the intermodal matching is not sufficient for the experience of ownership [Tsakiris, 2010]. In particular, it has been revealed that the RHI strictly requires physical, anatomical, postural and spatial plausibility of the real and fake hands [Tsakiris and Haggard, 2005, Costantini and Haggard, 2007, Lloyd, 2007, Makin et al., 2008] (see also [Liepelt et al., 2017]). Hence, the bottom-up integration of multisensory inputs seems to be modulated by experience-driven predictive information, which allows for active comparison between the properties of the viewed (non)-corporeal object and the internal model of the body [Tsakiris et al., 2008, Apps and Tsakiris, 2014]. The finding of Ferri and colleagues [Ferri et al., 2013, Ferri et al., 2017] further supported the fundamental role of the top-down processes in the modulation of body ownership. The authors demonstrated that the experience of ownership over a non-bodily object could originate as a consequence of pure expectation and anticipation of correlated exafference in the absence of actual tactile stimulation [Ferri et al., 2013, Ferri et al., 2017]. Together, this evidence supports the hypothesis that in the context of externally generated inputs (classical RHI), body ownership relies on two intertwined processes. Namely, (1) the bottom-up accumulation and integration of tactile and visual cues, and (2) top-down comparison between the novel sensory stimuli (i.e., rubber hand) and experience-driven priors about the internal model of the body [Tsakiris and Haggard, 2005, Blanke, 2012, Apps and Tsakiris, 2014, Clark, 2013, Seth, 2013]. We will refer to tactile or proprioceptive modalities as *proximal*, requiring an object to enter in direct contact with the surface of the body, and to the visual or auditory modalities as *distal*, sensing from a distance without getting in direct contact with the body.

Only recently the principles of body ownership have been studied in the context of self-generated (reafferent) sensory signals using physical set-ups (i.e., moving Rubber Hand Illusion, mRHI) [Dummer et al., 2009, Walsh et al., 2011, Ma and Hommel, 2015b, Tsakiris et al., 2006b, Kammers et al., 2009, Newport et al., 2010, Ma and Hommel, 2015a], or virtual reality (i.e., moving Virtual Hand Illusion, VHI) [Sanchez-Vives et al., 2010, Kalckert and Ehrsson, 2012, Shibuya et al., 2018]. In these protocols which include movement (mRHI and VHI), subjects are typically instructed to reach a specific target (goal-oriented) or to move the fingers/hand/arm continuously within a specific area (free exploration) while observing the (a)synchronously moving rubber or virtual analog. The results yield that there is a strong experience of ownership in the condition where the movements of the real and fake arms are spatiotemporally aligned [Dummer et al., 2009]. Contrarily, participants report no ownership of the fake body-part when the visual feedback of self-initiated movement is (inconsistently) delayed or displaced, and therefore does match the proprioceptive information [Blakemore et al., 2000]. Hence, similar to the classical RHI, in the context of self-generated movements, ownership seems to depend on the consistency of sensory information from proximodistal modalities, in this case, proprioceptive (proximal) and visual (distal). Interestingly, different to the classical RHI, in VHI as well as mRHI the experience of ownership emerges independently of whether (1) the visual, anatomical or structural properties of the avatar satisfy well-established priors about the own body [Banakou et al., 2013, Peck et al., 2013, Romano et al., 2015, Van Dam and Stephens, 2018, Ma and Hommel, 2015a], (2) there is a (consistent) delay in the visual feedback of the movement (3) the viewed object is ‘connected’ to participants’ body [Ma and Hommel, 2015a]. Crucially, the condition which needs to be satisfied is that the action-driven sensory feedback from proximodistal modalities matches the predicted one [Dummer et al., 2009, Sanchez-Vives et al., 2010, Ma and Hommel, 2015b]. In line with physiological and motor control studies [Proske and Gandevia, 2012, Miall and Wolpert, 1996, Wolpert and Flanagan, 2001], this evidence suggests that when moving in a goal-oriented manner body ownership is weighted stronger by the congruency of the internal (forward) model of the action and the action effects, the same mechanism which impacts agency [Hommel, 2009, Gallagher, 2007, Longo and Haggard, 2009, D’Angelo et al., 2018], rather than the (generative) model of the body and its physical specifics [Ma and Hommel, 2015b]. Crucially, it has been well established that the forward models are not limited to the bodily (proximal or proximodistal) feedback exclusively, but instead, they integrate across all sensory predictions which pertain to the interactions of an agent within an environment, including purely distal cues [Miall and Wolpert, 1996, Jordan and Rumelhart, 1992]. For instance, under normal conditions, the visuoauditory signals of the puck hitting the goal are spatiotemporally aligned with its trajectory that depends on the direction of the arm movement. However, if the actual location of the sound of the puck hitting the goal does not correspond to the efference copy or corollary discharge, it would reflect on the Sensory Prediction Errors (SPE) of the forward model [Wolpert et al., 1995, Miall and Wolpert, 1996, Wolpert et al., 2011, Woodgate et al., 2015, Maffei et al., 2017]. Thus if body ownership results from a consistency of forward models, it would be affected by the (in)congruency of not only proximodistal cues such as in the mRHI and VHI [Dummer et al., 2009, Sanchez-Vives et al., 2010] but also purely distal signals given that they constitute task-relevant information about the action-consequences.

Here, we propose that in contexts where the sensory signals are self-generated, such as in the moving Rubber Hand Illusion or the Virtual

Hand Illusion, body ownership depends on the sensory prediction errors from purely distal multisensory modalities, which would suggest a mechanism similar to the forward model or corollary discharge. We, therefore, hypothesize that the experience of ownership over a virtual body will be compromised when action-driven and task-relevant visuoauditory feedback of goal-oriented movements will not match sensory predictions. We also expect that the incongruency of those cues will affect performance. To test this hypothesis, we devise an embodied virtual reality-based goal-oriented task where action outcomes are signaled by distinct auditory signals. We manipulate the cues with respect to their spatial, temporal and semantic congruency, and compare body ownership and performance across two experimental conditions, where purely distal cues are either congruent or incongruent. Our results demonstrate, for the first time, that purely distal signals which violate predictions about the consequences of action-driven outcomes affect both performance and body ownership.

3.2 Materials and Methods

3.2.1 Participants

After providing written informed consent, sixteen healthy participants were recruited for the study, eight males (mean age 24.0 ± 2.65) and eight females (mean age 22.64 ± 2.25). Since no previous study assessed the effects of the congruency of purely distal modalities on body ownership, we could not perform a power analysis to determine the sample size. We, therefore, based the choice of *N* on previous studies [Mohler et al., 2010]. All subjects were right-handed (handedness assessed using the Edinburgh Handedness Inventory) [Oldfield, 1971], had normal or corrected-to-normal vision and reported normal hearing. They were pseudorandomly assigned to two experimental groups following a between-subjects design, which prevented habituation to the ownership measures, visuoauditory manipulations, and fatigue. We used stratified randomization to balance the conditions in terms of age, gender and previous experience with virtual reality. All participants were blind to the purpose of the study.

The experimental procedures were previously approved by the ethical committee of the University of Pompeu Fabra (Barcelona, Spain).

3.2.2 Task: Virtual Reality-Based Air Hockey Game

The experimental setup (Figure 3.1 A) comprised a personal computer, a motion detection system (Kinect, Microsoft, Seattle), a Head Mounted Display (HTC Vive, www.vive.com) and headphones. Similar to others [Sanchez-Vives et al., 2010, Grechuta et al., 2017b], here we used virtual reality as a tool to study the modulation of body ownership. The protocol was integrated within the virtual environment of the Rehabilitation Gaming System [Cameirão et al., 2010, Grechuta et al., 2016b]. During the experiment, while seated at a table, participants were required to complete a goal-oriented task that consisted in hitting a virtual puck into the goal (air hockey, Figure 3.1 A, B1). The virtual body was spatially aligned to the real body. Throughout the experiment, the participants’ arm movements were continuously tracked and mapped onto the avatar’s arms, such that the subjects interacted with the virtual environment by making planar, horizontal movements over a tabletop (Figure 3.1 A, B). To prevent repetitive gestures, at the beginning of every trial the puck pseudorandomly appeared in one of the three starting positions (left, center, right) (Figure 3.1 B2). The frequency of appearance of every starting position was uniformly distributed within every experimental session. Participants received instructions to place their hand in an indicated starting position and to execute the movement to hit the puck when its color changed to green (“go” signal). Each trial consisted of one “hit” which could end in either a success (the puck enters the goal) or a failure (the puck hits one of the three walls). At the end of every trial, participants were to place their left hand back at the starting position. The experimental block, in both conditions, consisted of 150 trials preceded by 20 trials of training (training block) (Figure 3.1 D) and followed by a threatening event. The threatening event served to measure autonomous responses to an unexpected threat (body ownership measure, Figure 3.1 C) [Armel and Ramachandran, 2003]. Overall, the task had an approximate duration of 20 minutes.

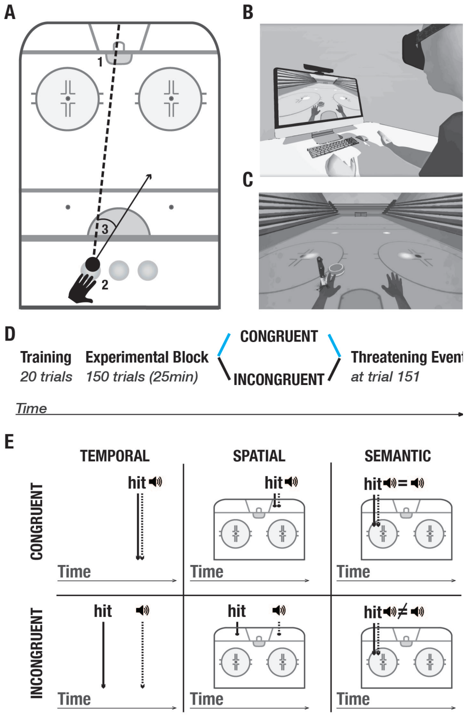


Figure 3.1: Experimental setup and protocol. (A) Task. 1- goal. 2- three starting positions. 3- example of a directional error, calculated as the difference between the actual direction vector and a straight line between the position of the puck and the goal. (B) Experimental setup. (C) Threatening event. (D) Experimental protocol. All participants underwent the training block. In the experimental block, they were randomly split into two conditions: Congruent (blue), and incongruent (black). At trial 151, all participants went through the threatening event which served to measure galvanic skin responses. The same color-code (congruent- blue, incongruent- black) is used in the following figures. (E) Purely distal visuoauditory manipulations- temporal, spatial and semantic. Upper panel: congruent condition; lower panel: incongruent condition.

3.2.3 Multisensory Feedback

Task-Relevant Visuomotor Signals. Throughout the experiment, participants were exposed to the visual feedback of self-generated arm movements. Specifically, the real arms were tracked by the motion detection input device and mapped onto the avatar's arms in real time allowing synchronous feedback. This method served to control for the congruency of proximodistal (i.e., visuomotor) signals which has been shown to underlie body ownership and agency [Sanchez-Vives et al., 2010]. It also guaranteed that the only manipulated variables were the distal modalities (i.e., visual and auditory).

Task-Relevant Visuoauditory Signals The task included task-relevant distal cues in the form of auditory feedback which was triggered as a consequence of every interaction of the puck with the environment. In particular, at the end of every trial, an auditory cue constituted a binary reinforcement signal informing about a failure (negative sound) or a success (positive sound). To study whether purely distal cues influence body ownership and performance, we manipulated the congruency of the auditory stimuli in three domains (Figure 3.1 E) — temporal: the time of the cue was synchronized with the time of the hit; spatial: the cue originated from the location of the hit, and semantic: the feedback of the cue reflected performance in a binary way (i.e., success or failure). The auditory cues were manipulated in two experimental conditions including congruent and incongruent. In the training block and the congruent condition, auditory cues were always congruent such that they occurred at the time of the hit, at the location of the hit, and they reflected performance. In the incongruent condition, the auditory signals were always incongruent. Namely, (1) the sound of the hit was anticipated or delayed, that is, it occurred randomly within 200-500ms before or after the actual collision (temporal domain), (2) it originated in a different location than the actual hit, that is, 5-15deg away from the actual hit, or (3) it did not reflect performance, that is participants heard the sound of failure following a successful trial and vice versa (semantic domain).

We chose those three manipulations to include all the dimensions necessary for the performance of the present task: direction, force, as well as the knowledge of results. Each of the dimensions (spatial, temporal, and semantic) provides unique information to the subject about the consequences of one's actions. Specifically, (1) the spatial dimension informs about the direction of the ballistic movement (where the puck hits the wall/goal), (2) the temporal dimension informs about the force applied to the action (when the puck reaches the wall/goal), whereas (3) the semantic dimension informs about the outcome of the action (either success or a failure). As such, all these dimensions contribute to the generation of prediction errors that can be integrated by an internal model to adjust motor performance. Spatial and temporal dimensions provide information about the action parameters on a continuous range and can be used as a supervising signal whereas the semantic dimension constitutes a binary reinforcement signal. All manipulations were pseudorandomly distributed and counterbalanced within each session to counteract order effects. Importantly, task-relevant proximodistal cues such as the visual feedback of the arm movements remained congruent in both conditions.

3.2.4 Measures

Motor control

We used three measures to quantify performance: scores, directional error, and reaction times. Scores were calculated as the percentage of successful trials (the puck enters the gate), while the directional error indicated the absolute angular deviation from the straight line between the starting position of the puck (left, central or right) and the center of the gate (Figure 3.1 B3). We computed the reaction times as time intervals between the appearance of the puck and action initiation. Since the task did not impose a time limit, we expected neither significant differences in reaction times between the conditions nor speed-accuracy trade-offs. We predicted that the manipulations of purely distal (visuoauditory) action-driven signals in the incongruent condition might alter scores and directional accuracy as compared to the congruent condition.

Body Ownership

Galvanic Skin Response (GSR). At the end of every experimental session, in both conditions, we introduced a threatening event (a knife falling to stab the palm of the virtual hand, Figure 3.1 C) to quantify autonomous, physiological responses to an unexpected threat [Armel and Ramachandran, 2003]. To prevent movement-driven muscular artifacts, we recorded the skin conductance responses from the right hand which did not move during the experiment. For the analysis, we calculated the mean and the standard deviation of the integral of the baseline (10 seconds time window before the threatening stimulus onset)-subtracted signal per condition in a non-overlapping time windows of 9s [Petkova and Ehrsson, 2008]. In particular, we expected an increase in the GSRs following the threatening stimulus in the congruent as compared to the incongruent condition.

Proprioceptive drift. Prior to and upon completion of the experiment, all the subjects completed the proprioceptive drift test which followed a standard technique, see for instance [Sanchez-Vives et al., 2010]. Specifically, the participants were asked to point to the location of the tip of their left index finger with the right index finger with no visual feedback available. The error in pointing [Tsakiris and Haggard, 2005] was computed as the distance between the two locations (the actual location of the tip of the left index finger and the pointing location) and measured in centimeters. We subtracted baseline responses from post-experimental errors for each participant. We expected stronger proprioceptive recalibration, and therefore, higher pointing errors in the congruent as compared to the incongruent condition.

Self-reports. At the end of every session, all participants completed a questionnaire which evaluated the subjective perception of body ownership and agency, adapted from a previous study [Kalckert and Ehrsson, 2012]. The entire questionnaire consisted of twelve items (Table 3.1), six per domain (ownership and agency), three of which were related to the experience of ownership and agency respectively, while the remaining served as controls. Participants answered each statement on a 7-point Likert Scale ranging from ‘-3’: being in strong disagreement to ‘3’: being in strong agreement.

To counteract order effects, the sequence of the questions was randomized across all the subjects.

Table 3.1: The questionnaire, consisting of 12 statements divided into four different categories.

Category	**Question**
Ownership	I felt as if I was looking at my own hand
	I felt as if the virtual hand was part of my body
	I felt the virtual hand was my hand
Ownership Control	It seemed as if I had more than one left hand
	It appeared as if the virtual hand were drifting towards my real hand
	It felt as if I had no longer a left hand, as if my left hand had disappeared
Agency	The virtual hand moved just like I wanted it to, as if it was obeying my will
	I felt as if I was controlling the movements of the virtual hand
	I felt as if I was causing the movement I saw, and the control questions were
Agency Control	I felt as if the virtual hand was controlling my will
	I felt as if the virtual hand was controlling my movements
	I could sense the movement from somewhere between my real and virtual hand

3.3 Results

To test our hypothesis that action-driven purely distal cues which pertain to the task contribute to body ownership, we used a virtual reality-based experimental setup (Figure 3.1 A, B) where subjects were to complete a goal-oriented task, and manipulated the congruency of auditory action outcomes (Figure 3.1 E). The experimental protocol (Figure 3.1 D) consisted of three phases: the training block, (2) the experimental block in either congruent or incongruent condition, and (3) the threatening event (Figure 3.1 D, C). To quantify body ownership, for each experimental session, we measured proprioceptive drifts, recorded Galvanic Skin Responses (GSR) to an unexpected threat, and administered self-reports. To measure performance, we computed scores, directional errors, and reaction times. For the analysis, we used t-tests and calculated Cohen’s d to evaluate differences between conditions and the associated effect sizes.

3.3.1 Motor Control

Firstly, our results showed that the normalized performance-scores (proportion of successful trials) were significantly higher in the congruent ($\mu = 0.35, sd = 0.47$) than in the incongruent condition ($\mu = 0.17, sd = 0.38$), ($t(14) = 8.89, p < 0.001, d = 0.42$) (Figure 3.2 A). To explore the effects of the congruency of purely distal signals on performance, we compared both conditions in terms of directional errors (Figure 3.2 B). In particular, a T-test indicated that the errors were significantly higher in the incongruent ($\mu = 6.42, sd = 4.52$) than in the congruent condition ($\mu = 3.30, sd = 2.01$), ($t(14) = 19.52, p < 0.001, d = 0.89$) (Figure 3.2 C). To further investigate the relationship between the quality of the distal cues and performance, we averaged and compared the directional errors following the three types of auditory manipulations (Figure 3.2 D). This analysis was performed exclusively for the incongruent condition. Interestingly, we found no difference between the distinct auditory cues including spatial ($\mu = 10.17, sd = 13.33$), temporal ($\mu = 7.99, sd = 9.75$) and semantic ($\mu = 7.22, sd = 7.23$) cues (Figure 3.2 D). Specifically, a Kruskal-Wallis test indicated that all manipulations had the same significant effect on body ownership ($x^2(2) = 1.74, p = 0.39$). In addition, we observed that the congruency of the distal cues had no significant effect on the averaged reaction times when comparing the incongruent group ($\mu = 0.48, sd = 0.05$) with the congruent group ($\mu = 0.51, sd = 0.01$), $p = 0.46$ (Figure 3.2 E).

3.3.2 Body Ownership

Prior to the appearance of the knife (10s baseline), the skin conductance was not different between the two groups ($t(14) = 0.60, p = 0.55$; $\mu =$

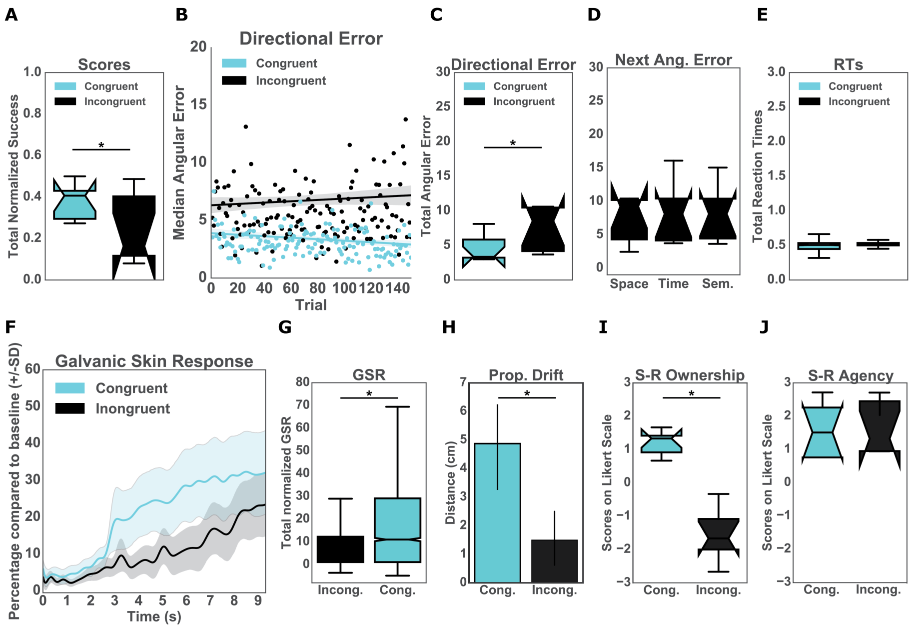


Figure 3.2: Upper panel: Performance. (A) Normalized percentage of successful trials per group. (B) Median directional error per trial over the experimental block (N=150) split per condition (C) Total directional error from all the trials per subject per condition. (D) This graph represents the mean values for the incongruent group only. In particular, the effects of the three auditory manipulations (spatial, temporal and semantic) on the mean directional error on the consecutive trials. (E) Mean reaction times from all trials per condition. *Lower panel: Body Ownership.* (F) Galvanic Skin Response (GSR). The sampling rate for the GSR signal was 60 Hz. Accordingly, the data was run through a low-pass filter with a cut-off frequency of 3 Hz. The plot represents the mean GSR and the associated standard deviation for all participants in a time window of 9s [Hägni et al., 2008], split per condition. The threatening event happened at time 0. (G) Mean GSR from 9s seconds post threatening event. (H) Proprioceptive drift. Results of the difference between pre- and post test calculated in centimeters per condition. (I) Score from the self-reported experience of body ownership per group. Scores above 0 indicate ownership. (J) Score from the self-reported experience of agency per group. Scores above 0 indicate the experience of agency.

$181.12, sd = 112.43$ for the congruent condition and $\mu = 230.25, sd = 183.75$ for the incongruent condition). The analysis revealed, however, that the post-threatening stimulus GSR was significantly higher in the congruent ($\mu = 42.54, sd = 33.98$) than in the incongruent group ($\mu = 29.67, sd = 26.82$) $t(14) = 21.03, p < 0.001, d = 0.42$ (Figure 3.2 F, G). Similarly, we found a difference in the proprioceptive drift between the congruent ($\mu = 4.88, sd = 2.36$) and incongruent group ($\mu = 1.5, sd = 1.51$) such that the errors in were significantly higher in the congruent condition ($t(14) = 3.4, p = 0.004, d = 1.7$) (Figure 3.2 H). We further report a statistically significant difference in the self-reported experience of ownership between the two conditions ($t(14) = 4.97, p < 0.001, d = 2.5$). The ownership ratings in the congruent group ($\mu = 1.13, sd = 0.56$) were greater than in the incongruent group ($\mu = -1.3, sd = 1.25$). We found no difference between the congruent ($\mu = -1.33, sd = 1.46$) and the incongruent group ($\mu = -1.3, sd = 1.25$) for the three control items ($t(14) = 1.79, p = 1.38$). We later analyzed questions related to agency. The results showed differences neither for the control questions ($t(14) = 0.22, p = 0.82$) between congruent ($\mu = -1.67, sd = 1.49$) and incongruent condition ($\mu = -1.83, sd = 1.48$) nor for the experimental ones, congruent ($\mu = 1.5, sd = 1.13$) and incongruent condition ($\mu = 1.33, sd = 1.48$). In both groups participants experienced high agency during the experiment.

3.3.3 Relationship of the Ownership Measures

We assessed the relationship between the objective, subjective and behavioral ownership measures and, per each participant in both conditions, we computed: (1) mean GSR from nine seconds post-threat, (2) mean of the three ownership questions; and (3) baseline-subtracted proprioceptive drift. The Spearman rank-order correlation between post-threat GSR and self-reported ownership was close to significance ($r = 0.47; p = 0.06$) (Figure 3.3 A). However, we report high and significant positive correlation between the proprioceptive drift and self-reported ownership ($r = 0.75; p < 0.001$) (Figure 3.3 B) as well as between the post-

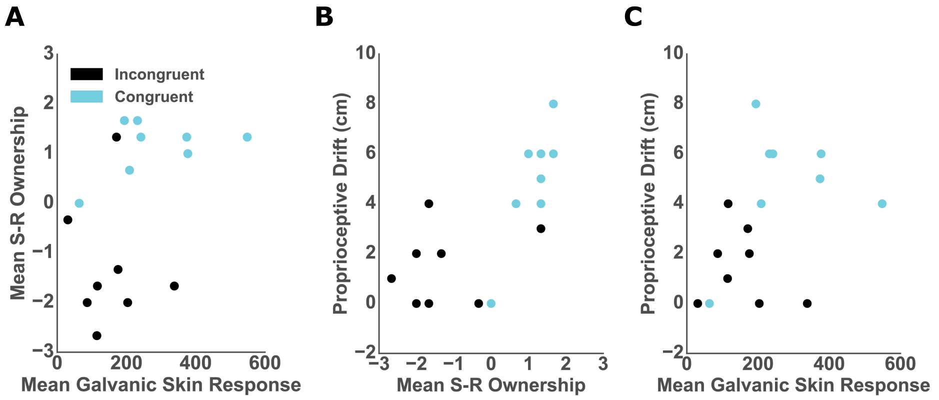


Figure 3.3: *Correlations* In all graphs dots represent individual participants and colors represent conditions: blue- congruent and black- incongruent. (A) Mean GSR 9s post-threatening event and mean self-reported ownership. (B) Mean self-reported ownership and the proprioceptive drift score. (C) Mean GSR 9s post-threatening event and and the proprioceptive drift score.

threatening GSR and proprioceptive drift ($r = 0.52; p < 0.03$) (Figure 3.3 C).

3.4 Discussion

In this study, we asked whether body ownership depends on the consistency of task-relevant purely distal sensory cues which result from self-initiated actions. In particular, we investigated the influence of those cues on performance and ownership using an embodied, virtual reality-based goal-oriented task where action outcomes were signaled by distinct auditory signals. We manipulated the congruency and therefore the predictability of those reafferent sensory signals and hypothesized that the (in)congruency of visuoauditory stimuli would affect both performance and body ownership. Our results support our prediction and suggest that both are compromised when action-driven purely distal signals are incongruent.

The plasticity of body ownership relative to the spatiotemporal coincidence of exafferent and reafferent multisensory signals has been well-accepted [Botvinick and Cohen, 1998, Craig, 2002, Tsakiris, 2010, Seth, 2013, Suzuki et al., 2013, Blanke, 2012]. In particular, neurophysiological and behavioral studies have demonstrated that the experience of ownership is established through bottom-up integration and top-down prediction of proximodistal cues within the peripersonal space [Rizzolatti et al., 1981a, Makin et al., 2007, Tsakiris, 2010, Blanke, 2012]. Crucially, however, depending on whether the sensory signals are externally (classical Rubber Hand Illusion, RHI) or self-generated (*moving* Rubber Hand Illusion, mRHI and moving Virtual Hand Illusion, VHI), the top-down expectation signals seem to be qualitatively different. On the one hand, in the RHI, the sensory correlations are modulated by top-down influences which constitute empirically induced priors related to the internal model of the body [Tsakiris and Haggard, 2005, Costantini and Haggard, 2007, Lloyd, 2007, Makin et al., 2008]. For instance, the illusion of ownership will not occur if the shape or the location of the fake hand is not plausible [Tsakiris, 2010]. On the other hand, the evidence from the mRHI and VHI supports that, in the contexts of self-generated stimuli, body ownership is actively shaped by top-down processes allowing for continuous comparison between the actual and predicted action-consequences from proximodistal modalities [Dummer et al., 2009, Sanchez-Vives et al., 2010, Ma and Hommel, 2015b, Ma and Hommel, 2015a]. In fact, when the errors in those sensory predictions (the so-called Sensory Prediction Errors, SPE) are insignificant, that is, when the visual feedback of the position of the rubber (mRHI) or virtual (VHI) hand is congruent with the proprioceptive cues, the ownership over the artificial arm is high, and vice versa [Dummer et al., 2009, Sanchez-Vives et al., 2010]. Contrary to the standard RHI, in the mRHI and VHI, the physical, spatial and temporal characteristics of the body do not influence the experience of ownership [Banakou et al., 2013, Peck et al., 2013, Romano et al., 2015, Van Dam and Stephens, 2018, Ma and Hommel, 2015b]. Moreover, it has been demonstrated that participants can perceive an actively operated virtual non-corporeal and 'disconnected' object (balloon or a square) as an extension of their own body as long as it follows the predicted trajectory [Ma and Hommel, 2015a]. Thus, when acting in the world, the top-down predictive processing modulating ownership seems not to depend on the generative models of self but rather on the forward models (or corollary discharge) which guide action by generating sensory predictions about the consequences of movement based on the efference copy [Miall and Wolpert, 1996, Kilteni and Ehrsson, 2017, Sanchez-Vives et al., 2010, Ma and Hommel, 2015b]. Similar, from the perspective of ideomotor theory, ownership might be viewed as depending on the difference between the goals (intended action effects) and the perceptual consequences (actual action effects) [Stock and Stock, 2004, Hommel, 2009, Shin et al., 2010].

Ample research demonstrates that the central nervous system uses forward models for the differential processing of the predicted and the actual reafferent information which was shown to underlie motor control and agency [Miall and Wolpert, 1996, Wolpert et al., 1995, Bäß et al., 2008, Sommer and Wurtz, 2008, Crapse and Sommer, 2008, Schwarz et al., 2018]. Crucially, the internal (forward) models do not exclusively process sensory signals related to the body, but they integrate across all sensory information from both proximal (proprioceptive, tactile) and distal (visual, auditory) modalities [Miall and Wolpert, 1996, Jordan and Rumelhart, 1992]. This would suggest that, if body ownership results from a consistency of forward models, it will be affected by the(in)congruency of not only proximodistal cues such as in the moving rubber hand illusion or the virtual hand illusion [Dummer et al., 2009, Sanchez-Vives et al., 2010] but also purely distal signals given that they constitute information about the action-consequences. In this study, we explicitly addressed this question using a variation of a VHI paradigm, which required the participants to perform actions that triggered distal (auditory) cues. Those auditory cues indicated the location and the time of a collision of a puck with the walls or the goal as well as the outcome (failure or success). To test whether action-driven and task-relevant sensory signals impact body ownership, in one of the groups, we manipulated their congruency. We predicted that the ownership scores, measured subjectively, objectively and behaviorally could be lower in the condition where the cues do not match predictions about purely distal sensory signals.

Did the proposed purely distal cues affect body ownership? Results from all the ownership measures (Figure 3.2 Lower panel: Body Ownership), including skin conductance (GSR), proprioceptive drift and the questionnaire support that purely distal cues which pertain to the task and violate predictions about the auditory action outcomes compromise body ownership. Specifically, we found that the scores were significantly higher in the congruent compared to the incongruent condition in all analyses (Figure 3.2 Lower panel: Body Ownership). Subsequent correlations between the proposed measures (Figure 3.3) further confirmed the consistency of the obtained results within three dimensions of ownership quantification including physiological response, behavioral proprioceptive recalibration, and a conscious report [Longo et al., 2008]. Similar to the mRHI, VHI [Sanchez-Vives et al., 2010, Ma and Hommel, 2015b] and their variations (i.e., [Ma and Hommel, 2015a]), here we interpret the obtained low-ownership outcome in the incongruent condition (Figure 3.2 Upper panel: body Ownership) as a consequence of high sensory prediction errors possibly computed but the forward model [Miall and Wolpert, 1996, Crapse and Sommer, 2008, Apps and Tsakiris, 2014, Limanowski and Blankenburg, 2013]. In our case, however, the sensory conflicts were driven by a discrepancy between the predicted and actual purely distal visuoauditory signals which did not pertain to the body but were relevant to the outcome of the goal-oriented task. We speculate that the manipulation of the proposed signals might have reflected on the errors of the forward models which influence performance and possibly body ownership [Wolpert et al., 1995]. This could further suggest that the integration of signals from distal modalities might affect the integration of signals from proximal or proximodistal modalities establishing a feedback loop. In such case, any (in)congruent relationship between distal, proximodistal, and proximal signals which pertain to the goal of the task would affect the experience of ownership and even define the boundaries of the embodied self. To the best of our knowledge, our results propose for the first time that the ownership of a body might be driven by bottom-up integration and top-down prediction of purely distal modalities occurring outside of the body and outside of the peripersonal space [Rizzolatti et al., 1981a]. This would support recent findings which suggest that body ownership is coupled to the motor systems and that, similar to the experience of agency, it might depend on the congruency of a forward model or corollary discharge [Grechuta et al., 2017b, Kilteni and Ehrsson, 2017, Ma and Hommel, 2015b]. As expected, the visuoauditory manipulations did not significantly influence the perceived agency (Figure 3.2 J). Participants reported control over the virtual hand in both conditions, probably due to the congruent mapping of the proximal cues (see Methods section about the sensory manipulations). The visual feedback of the movement of the arm always followed the desired trajectory, which is one of the three questions addressed in the standard self-reported agency assessment [Kalckert and Ehrsson, 2012].

At the current stage, two questions remain open. First, how can the integration of distal and proximodistal cues occur in the service of body ownership? Since the primary purpose of the present study was to investigate the influence of purely distal signals on body ownership, the proximodistal (visuo-proprioceptive) cues within the peripersonal space were congruent in both groups. Indeed, based on those cues, participants could always predict the location and the time of the distal auditory signals (spatial and temporal manipulation) as well as the outcome of an action (semantic manipulation). Therefore, in the incongruent condition, where the distal consequences of the actions did not match the predictions, we expected that the sensory prediction errors would negatively impact ownership. However, with the current design, we can neither explain the interaction of the proximodistal and distal cues nor how do they weight the experience of ownership. Future studies should further investigate the relationship between the visual and auditory cues and their relative impact on body ownership by, for instance, manipulating visuomotor and visuoauditory feedback independently during a motor task. A recent Hierarchical Sensory Predictive Control (HSPC) theory proposes a cascade of purely sensory predictions which mirror the causal sequence of the perceptual events preceding a sensory event [Maffei et al., 2017]. In the context of anticipatory control, this control architecture acquires internal models of the environment and the body through a hierarchy of sensory predictions from visual (distal) to proprioceptive and vestibular modalities (proximal). If body ownership and motor control share the same forward models, which comprise both distal, proximodistal, and proximal signals, ownership might be realized through a similar cascade of sensory predictions. In our case, however, which includes a goal-oriented task and voluntary control, the internal models might be acquired from the proprioceptive and vestibular modalities (proximal) to visual (distal), a hypothesis yet to be investigated. In such case, one could expect differences in reaction times between the congruent and the incongruent conditions due to increased sensory prediction errors. Interestingly, our results yielded no differences in the reaction times between the groups. We believe that this result might depend on the congruency of proximodistal signals. Specifically, the visual feedback of the movement always matched the proprioceptive cues. It is possible that for motor control the prediction errors from the proximodistal modalities are more relevant (they are weighted higher) than those from purely distal. We suggest that future studies should systematically investigate the contribution of different cues to performance, possibly within the framework of HSPC [Maffei et al., 2017]. Second, if body ownership depends on the consistency of internal models, and therefore on the accuracy of sensory predictions, could task-irrelevant signals manipulate it? While playing Air Hockey, the brain does not only integrate action-driven sensory signals but also simultaneously processes purely external action-independent information which derives from the environment. This information might well include corrective information and, therefore, be relevant to the task (i.e., the wind which affects the trajectory of the puck) or not (i.e., time of the day) [Shadmehr et al., 1994]. Changing the rules of the environment and investigating the experience of ownership and performance when action-independent (task-irrelevant) sensory expectations are violated would shed light on the nature of sensory signals relevant for the processing of self as well as their underlying mechanisms (i.e., generative and forward models) [Friston, 2012, Apps and Tsakiris, 2014, Seth, 2013].

What is the role of purely distal action-driven cues in goal-oriented behavior? Our results demonstrate that performance, as measured through the overall scores (Figure 3.2 A) and directional errors (Figure 3.2 B, C), was significantly hampered in the incongruent compared to the congruent condition. Importantly, these results did not depend on differences in reaction times (Figure 3.2 E) suggesting no influence of possible attentional biases (i.e., distractions) in either of the groups. On the one hand, this outcome might be interpreted within the framework of computational motor control. The reported differences in performance between the two conditions could have been influenced by the discrepancies between the efference copies of distal events and the actual action outcomes. Indeed, results from motor control studies support the notion that learning (progressive reduction of error) depends on both proximal and distal sensory prediction errors that allow for adjustments and anticipation of possible perturbations deriving from the body and environment [Jordan and Rumelhart, 1992, Mazzoni and Krakauer, 2006, Tseng et al., 2007, Krakauer, 2009, Morehead et al., 2017, Maffei et al., 2017]. As a result, during action execution, inputs from all the sensory modalities are transformed into error signals updating the forward model and, consequently, future behavior [Wolpert and Kawato, 1998, Kawato, 1999, Shadmehr et al., 2010, Wolpert et al., 2011]. In our experiment, the directionality of the error indicated by the spatial distribution of the sound, the speed of the puck indicated by the temporal characteristics of the sound, as well as the knowledge of results all constituted error signals which could supervise corrective motor commands. Crucially, while the spatial and temporal dimensions provided information about the action parameters on a continuous range, the semantic dimension constituted a binary reinforcement signal informing about a failure or a success. As such, the chosen audiovisual cues in the incongruent condition might have influenced performance, which, in turn, affected body ownership. In fact, clinical studies provide evidence that patients suffering from hemiparesis, whose motor function is reduced due to stroke, progressively stop using the paretic limb: the so-called learned non-use phenomenon [Taub et al., 2006]. In this, and other neurological cases, a prolonged lack of use (low performance) often causes disturbances in the sense of ownership and agency [Gallagher, 2006] supporting a hypothesis that there might be a causal effect between performance and body ownership. The present design which includes three types of sensory manipulations pseudorandomly distributed within each block does not allow us to disambiguate between the specific contribution of each of the manipulations. A systematic study on the influence of individual sensory signals, including the three manipulations, would help to better understand the mechanisms accounting for low-performance scores in the incongruent condition. An alternative interpretation of our results is related to the experimental and theoretical framework of body ownership. Several studies propose that body ownership is coupled to the motor system such that it updates the sensory representation of the body and provides inputs to the forward model. The forward model, in turn, generates and updates predictions relative to both the body and the environment during voluntary actions [Kilteni and Ehrsson, 2017], reinforcing the history of sensorimotor contingencies. In particular, we find evidence that body ownership is involved in generating body-specific predictions about the sensory consequences of voluntary actions thus determining somatosensory attenuation [Kilteni and Ehrsson, 2017]. This finding is consistent with another study which employed a standard RHI in virtual reality and demonstrated that ownership is correlated with motor performance during a perceptual decision-making task [Grechuta et al., 2017b]. Contrary to the previous discussion, in this case, ownership would have a modulatory effect on performance.

At the current stage, we cannot disambiguate between the two alternative hypotheses and determine whether the integration of purely distal cues influences ownership and performance in parallel or independently and what is the directionality. We demonstrate, however, that both depend on the congruency of action-driven and task-relevant purely distal signals, which supports the notion that both rely on the consistency of forward models driving goal-oriented action [Seth, 2013, Apps and Tsakiris, 2014]. We expect that this outcome will allow for the advancement of our understanding of the mechanisms underlying body ownership. To improve the experimental quality of the present study and further support our findings, future studies shall consider a bigger sample size as well as an alternative objective measure of ownership (i.e., body temperature) which would allow for conducting a within-group experiment without biasing the physiological signals [Moseley et al., 2008]. Finally, the reported finding might find applications in fields such as motor training simulators and rehabilitation. For instance, virtual reality-based treatments of post-stroke motor disorders [Cameirão et al., 2010, Grechuta et al., 2016b, Ballester et al., 2015b, Mihelj et al., 2014] might benefit from a design of reliable and spatiotemporally congruent environments which may positively impact the ownership of the virtual body as well as performance possibly impacting recovery. Further clinical studies should evaluate the same principle in rehabilitation protocols for ownership disturbances following acquired brain lesions including neglect [Coslett, 1998], anosognosia for hemiplegia [Pia et al., 2004] or somatoparaphrenia [Fotopoulou et al., 2011].

Chapter 4

(BODY) OWNERSHIP AS A MATTER OF CONTROL, NOT PRIORS, CONNECTEDNESS OR PERIPERSONAL SPACE

This chapter is based on:
De la Torre, J., Grechuta, K., Daversa, D., & Verschure, P. (2019) (Body) ownership as a matter of control, not priors, connectedness or peripersonal space. *In preparation.*

4.1 Introduction

While playing computer games, we usually have an avatar who we can fully control and who takes part in the game on our behalf. Interestingly, at the end of a successful round, we would typically shout “I won!!!” and not “Super Mario Bros won!!!” This happens independently of whether the agent is located in front of us on the screen or further away such as in the case of virtual reality. Interestingly, although the of identification with virtual objects, tools or avatars might not seem surprising but rather quite natural, it has not been given much attention in the field of cognitive psychology despite its curious implications for the understanding of the principles of body ownership. For instance, it raises a puzzling question of what are the constrains of body ownership in the context of active control?

From the evolutionary perspective, the ability to recognize one’s own body and distinguish it from other agents and objects in the environment is fundamental for survival. However, self-recognition is such an inherent feature in developed organisms that it is often taken for granted. Interestingly, using experimental methods such as bodily illusions, a number of studies have demonstrated that the so-called body ownership *can* be flexibly altered more than previously assumed. For instance, studies which explore mechanisms underlying the emergence of body ownership in the context of action (Virtual Hand Illusion, VHI) [Sanchez-Vives et al., 2010, Kalckert and Ehrsson, 2012, Shibuya et al., 2018, Grechuta et al., 2019d], converge to suggest that body ownership (1) does not require physical plausibility of the effector (i.e., the controlled body or object) [Banakou et al., 2013, Peck et al., 2013, Romano et al., 2015, Van Dam and Stephens, 2018, Ma and Hommel, 2015a, Short and Ward, 2009], (2) does not require the effector to be connected [Ma and Hommel, 2015a, Short and Ward, 2009], and (3) depends on all, not only proximodistal but also purely distal, action-driven cues which pertain to the task [Grechuta et al., 2019d].

For example, as discussed in Chapter 3, in the context of a goal-oriented game such as the air-hockey, subjects report a high feeling of body ownership of the match the sensory expectations (i.e., the movement of the virtual hand follows the executed command, the trajectory of the puck is in line with the hit, and the auditory feedback of the puck hitting the goal is spatially and temporarily congruent). Moreover, one can also feel ownership over a non-corporeal object disconnected from the body such as a cone [Short and Ward, 2009] or a balloon [Ma and Hommel, 2015a] provided that it is fully controllable, that is, that the (multisensory) behavior of this object matches the predicted sensory feedback, similar to when we use tools [Iriki et al., 1996, Maravita and Iriki, 2004, Farnè and Làdavas, 2000, Holmes et al., 2004].

On the one hand, this evidence supports the hypothesis that the experience of body ownership in the context of action could arise from the congruency between the actual sensory signals and the prediction of a forward model (or corollary discharge). Indeed, the role of a forward model is to anticipate sensory consequences of goal-oriented movements and compare them to the actual signals in the guidance of motor behavior [Miall and Wolpert, 1996, Wolpert and Flanagan, 2001, Crapse and Sommer, 2008]. On the other hand, those results challenge the notion of the peripersonal space as a necessary condition for the experience of ownership [Ehrsson, 2012]. Specifically, the discussed findings suggest that the feeling of ownership may be experienced over any controllable object, independently of its location.

In this study, we explicitly addressed this question and investigated whether controllability is a sufficient condition to induce the feeling of body ownership regardless of the proximity of the effector. To test this hypothesis, we used a virtual reality-based setup, where the participants learned to control a disconnected object (i.e., a log) located outside of their peripersonal space. We hypothesized that the experience of ownership over this log will be higher in the condition where the mapping between specific arm movements and the resulting movements of the log is congruent. Thus it would allow full prediction and control of its behavior as compared to the condition where the mapping is random such that the behavior cannot be predicted and there is no experience of control. Our results support the notion that the experience of body ownership is driven by *control*, not priors, connectedness or the peripersonal space.

4.2 Method

Participants

Thirty-two healthy subjects from Universitat Politècnica de Catalunya were recruited for the study: 17 males (Mean ± SD: 24.75 ± 6.51 years old) and 15 females (Mean ± SD: 22.86 ± 3.66 years old). The sample size was based on previous studies [Kilteni et al., 2018]. All participants were right-handed, reported normal hearing, and normal or corrected-to-normal vision. Participants provided their written consent, and they were pseudorandomly assigned to one of the two experimental conditions, following a between-subjects design. The groups were balanced regarding age, gender, and previous experience with virtual reality.

Experimental Setup and Virtual Environment

The setup (Figure 4.1 A) consisted of a personal computer, head-mounted display (HTC Vive, www.vive.com), HTC Vive hand-held controllers and active noise control headphones (Beats Electronics LLC, California, USA) which ensured isolation from external sounds. The virtual environment was developed using Unity3D (Unity Technologies, SF, Copenhagen, Denmark), and it was scaled 1:1 with the real-world such that each distance-unit in the virtual world corresponded to the same distance-unit in the physical world. During the experiment, participants were seated in a rotatory chair, which allowed them to visually explore all the virtual environment from a fixed position by rotating their head. The camera through which they saw the environment was located in the middle of the virtual scene (Figure 4.1 B) and its height was adjusted to each participant to match the height of their eyes.

Task

During the experiment, participants were immersed in a scenario which resembled a garden (Figure 4.1 B). The environment included different obstacles: (1) rocks which impeded the advance and had to be surrounded, (2) steps which could be jumped on, and (3) rivers that could be crossed using the bridges. The task was to control a virtual object (a log) and direct it to assigned target locations by performing five pre-mapped gestures that triggered different behaviors of the log. Each participant started the experiment from the same initial position (see Figure 4.1 D2). Importantly, participants did not have a virtual body, and the only element they were able to control was the log located outside of the peripersonal space (Figure 4.1 D).

Control

Participants could control the log via HTC controllers. By performing five specific arm movements, they could trigger five specific movements of the log, thus changing its position in the virtual environment. In particular, (1) a horizontal movement of the arm from the center towards the left made the log move one step forward, (2) a horizontal movement of the arm from the center towards the right made the log move one step backwards, (3) a vertical movement of the arm from the center upwards made the log jump one step forward, (4) a rotation of the wrist from the center towards the left made the log rotate to the left, (5) a rotation of the wrist from the center towards the right made the log rotate to the right (see Figure 4.1 C for the details of the mapping). All the movements were to be instantiated in the starting position (Figure 4.1 C). Participants had to indicate the beginning of each movement by pressing the controller trigger. The direction of the log was indicated by its color, in particular, the front face was white, and the back face was black (Figure 4.1 C).

Protocol

Each experimental session was divided into two parts which lasted 5 and 15 minutes, respectively (Figure 4.1 E). In the first part of the experiment, participants underwent the training block where they were instructed to freely explore the virtual space. Since initially subjects did not know the mapping between the specific arm movements and the log, this block served to get familiar with the virtual environment and to learn to control it. In the second part of the experiment, subjects underwent the experimental block. At the beginning of this block, they were presented with a target in the form of a fountain (Figure 4.1 B Lower panel). Their task was to direct the log to the target by moving it with the controllers as learned during the training block. Participants received instructions to reach the target using the shortest path. After every successful trial, the target disappeared, and a new one appeared in a different location. The locations of the fountains were randomly distributed and matched for each participant. The length of every experiment was twenty minutes including the training

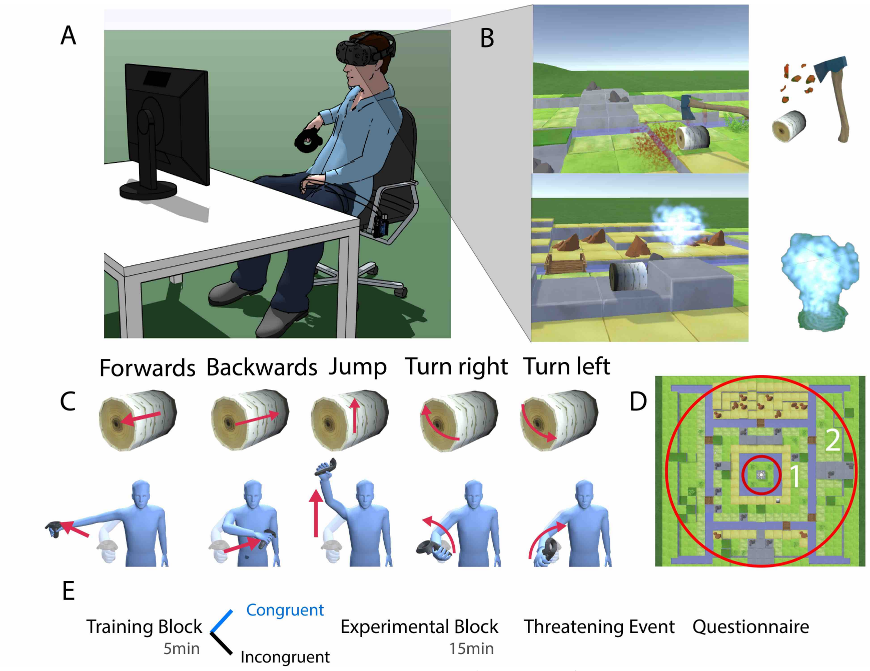


Figure 4.1: Experimental Setup and Protocol. **(A) Experimental setup.** The participant is seated on a chair in a fixed position and views the virtual environment through a head-mounted display. S/he controls the log using HTC controllers **(B) Participant's View.** Two examples of a first-person view of the virtual environment. Upper panel: threatening event (i.e., an axe stabbing the log), Lower panel- A targets (i.e., fountain).**(C) The mapping between the arm movements and the log.** The images represent five different controller movements associated with a five different log movement. In the congruent condition, the mapping was fixed, while in the incongruent condition the five movements triggered random responses. The central starting position is shown in transparent. **(D) Schematic bird view of the virtual environment.** The participant was positioned in the center of the map and could view the environment by rotating their head. Circle *1* marks the peri-personal space which could not be reached by the log, while circle *2* marks the extrapersonal space which is the control space of the log. **(E) Experimental protocol.** Participants are randomly split into the congruent or incongruent condition. During the training block, they freely explore the environment and learn the mapping implicit mapping with the log. During the experimental block, their task is to control the log and direct it towards the fountain-targets using the shortest path. At the end of the experiment, all the participants experience a threatening event. After the experiment, the participants are complete the questionnaires (see Methods, Measures).

block, therefore, the number of targets reached per participant depended on individual performance.

To study whether motor control over a non-corporeal object disconnected from the body and located outside of the peripersonal space is enough to induce ownership over this object, we manipulated the congruency of the mapping between the arm movements and the log in two conditions. As a result, subjects were split into two groups: congruent and incongruent (Figure 4.1 D). In the congruent condition, the mapping between the arm movements and the log was always the same such that each of the five gestures triggered that same behavior of the log as presented in Figure 4.1 C. Conversely, in the incongruent condition, the set of five gestures each time produced a different outcome, such that the consequences of self-generated movements did not have a constant effect on the log, and therefore, could neither be learned nor predicted. The incongruent condition served as control where we did not expect to find the experience of ownership and agency. We predicted, however, that the control over the log in the congruent condition might positively impact all self-reports.

In both conditions, at the end of the experimental block, participants experienced the so-called threatening event which consisted of an axe stabbing the log repeatedly (Figure 4.1 B, Upper panel). This block lasted 12 seconds during which physiological responses were recorded and stored to objectively quantify the experience of ownership [Armel and Ramachandran, 2003] (see section Measures).

Measures

Self-Report: First, to evaluate whether participants felt immersed in the proposed environment, we administered presence questionnaire based on [Witmer and Singer, 1998]. Secondly, to evaluate the subjective experience of ownership of the virtual log as well as the agency, subjects completed a 12-item questionnaire adapted from [Longo and Haggard, 2009, Kalckert and Ehrsson, 2012] where six questions served as controls. In addition to the standard items, we introduced two novel questions which had the goal to evaluate the perception of the boundaries of the self. These items were the following: "I felt as if the boundaries of my body expanded" and "I felt as if the log was an extension of me," respectively. We referred to this measure as the *Expanded Boundaries questionnaire*. In all questionnaires, the answers were to be delivered after the end of the experiment on a 7-Points Likert scale ranging from '-3': strongly disagree to '3': strongly agree. The order of the questions per each domain was randomized between subjects. The analysis of the questionnaires was performed by averaging the responses per participant and comparing them between the congruent and incongruent conditions.

Galvanic Skin Response (GSR): Similar to other studies, [Armel and Ramachandran, 2003, Sanchez-Vives et al., 2010], we measured GSR responses during the threatening event, in particular, when a virtual axe stabbed the log. To control for possible muscular artifacts, the GSR was recorded from the two fingers of the left hand which was still during the whole experiment. The signal was recorded using BITalino (https://bitalino.com/en/) at a sampling rate of 100Hz. In order to isolate phasic from the tonic response, we performed the Continuous Decomposition Analysis (CDA) [Benedek and Kaernbach, 2010a]. To quantify the increase of the GSR, for each group we subtracted the averaged first two seconds post-stimulus-onset from the late response (last two seconds).

4.3 Results

The present study aimed at determining whether full controllability and predictability of the behavior of a physically disconnected object located outside of the peripersonal space can induce body ownership over this object and bias the boundaries of the perceived self. To this end, we developed a virtual reality-based paradigm where participants did not owe a virtual (human-like) body but, instead, they had to control a virtual object, a log located outside of the peripersonal space, which responded accordingly to the movements of their right hand (see Figure 4.1 C for the mapping).

The protocol included training which served for the subjects to learn the mapping between the arm and the log. This phase was followed by a goal-oriented task in which the participants were required to direct the log towards virtual targets (fountains) by using the previously learned movement-patterns. To test our prediction, we manipulated the congruency of the mapping between the gestures and the movements of the log in two conditions: the congruent condition where the mapping was entirely predictable and the incongruent condition in which the mapping was unstable. In the congruent condition, the learned set of five motor commands always triggered the same movements of the log such that the behavior of the disconnected object was always fully predictable. However, in the incongruent condition, the movements of the log could neither be learned nor predicted. In both groups, we recorded Galvanic Skin Response (GSR) and collected self-reports which constituted physiological (i.e., objective) and conscious (i.e., subjective) measures of the experience of ownership, respectively. To ensure that the participants in both conditions felt equally immersed in the virtual environment, we administered presence questionnaire adapted from [Witmer, 1998].

4.3.1 Presence and Agency

First, we assessed the perceived experience of presence and agency (Figure 4.2 A and B) as reported in the questionnaires. In both conditions participants felt immersed in the virtual environment. Specifically, we found no statistically significant difference between the groups ($p = 0.21$; Congruent "C" group: $\mu = 1.12$, $std = 0.94$; Incongruent "I" group: $\mu = 0.59$, $std = 1.36$), which suggests that the incongruent mapping in "I" did not affect the subjective illusion of being surrounded by the artificial scenario. As expected, however, the analysis yielded a significant difference between the groups in the perceived experience of Agency ($p < 0.001$; "C" group: $\mu = 1.61$, $std = 1.46$; "I" group: $\mu = -1.47$, $std = 1.04$).

To assess whether the proposed mapping had an effect on the control of the log, we stored the times when the participants reached the target fountains, and computed the performance in both conditions. Indeed, we found a significant difference between groups ($p < 0.01$; "C" group: $\mu = 6.64$, $std = 4.28$; "I" group: $\mu = 0.16$, $std = 0.37$).

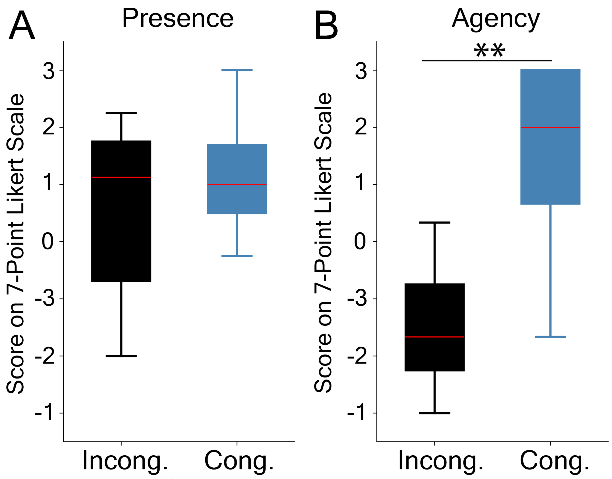


Figure 4.2: Self Reports. **Left.** Self-reported experience of Presence. **Right.** Self-reported experience of Agency. In both graphs, the values on Y-axis indicate responses on a 7-point Likert Scale.

4.3.2 Subjective and Objective Experience of Ownership

To quantify the subjective experience of ownership of the log we first analyzed self-reported answers. The analysis revealed statistical differences in the perceived experience of ownership between the two conditions ($p < 0.05$; "C" group: $\mu = -0.18$, $std = 1.32$; "I" group: $\mu = -1.21$, $std = 1.37$) (Figure 4.3 C Upper panel). Crucially, we found no differences in the control questions between groups ($p = 0.14$; Congruent "C" group: $\mu = -0.38$, $std = 1.49$; Incongruent "I" group: $\mu = -0.85$, $std = 1.11$). This result suggests that, indeed, the difference in controllability between experimental conditions could have a significant effect on the subjective experience of body ownership.

To corroborate this result and quantify the experience of ownership objectively, in both conditions we analyzed the Galvanic Skin Response

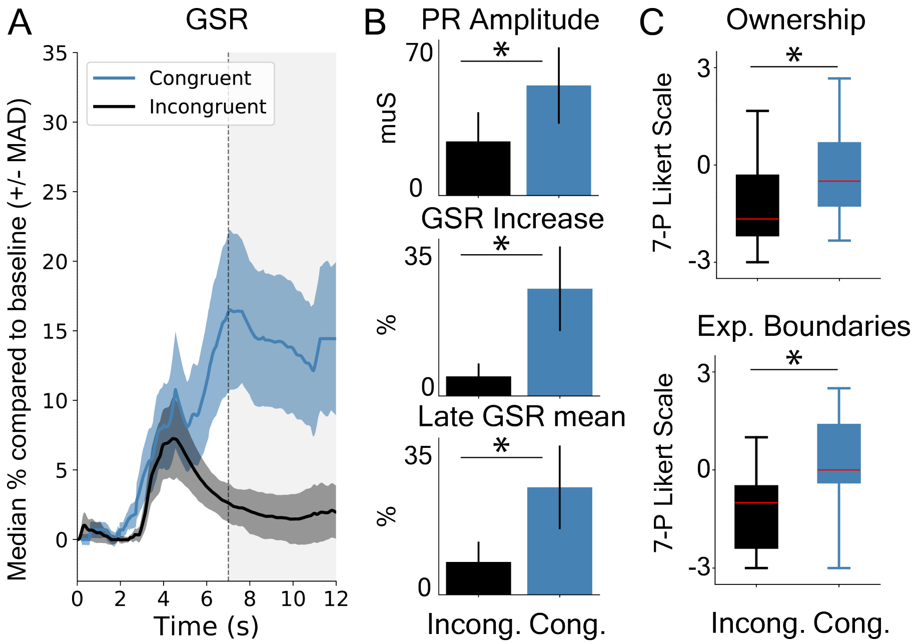


Figure 4.3: **(A)** Evolution of the GSR. Median (+/- MAD) responses per condition for all the participants in a time window of 12 seconds. The grey area represents the interest region of the late response (7-12s). Signals are normalized for each participant by subtracting the mean signal from 10 seconds prior to the stimulus-onset (the moment when the log is stabbed). **(B)** Top: GSR Phasic response amplitude (muS). Middle: Differences in the averaged increase of the GSR computed by subtracting the early response ($0s - 2s$) from the late response ($10s - 12s$). Bottom: Mean GSR during the late response ($7s - 12s$) compared to baseline per condition. **(C)** Top: Self-reported experience of ownership. Responses on a 7-point Likert Scale. Bottom: Self-reported score for the proposed "Extended-self questionnaire."

following the threatening event compared to baseline. Importantly, the analysis revealed a statistical difference between groups in the late response (7-12s) post-threat ($p = 0.05$; "C" group: $\mu = 24.38\%$, $std = 13.03$; "I" group: $\mu = 7.45$, $std = 15.24$) (Figure 4.3 B). The comparison of the increase between the groups yielded a statistically significant difference ($p = 0.03$; "C" group: $\mu = 25.59\%$, $std = 10.21$; "I" group: $\mu = 4.68$, $std = 10.07$). Subsequent testing revealed that the total amplitude in the phasic response was significantly higher in "C" than in "I" ($p =$

0.05). This result suggests that, also from an objective perspective (i.e. an autonomic response to a threat), body ownership is significantly higher in the congruent condition than in the incongruent one.

In addition to the significant differences in the standard ownership measures (i.e., questionnaire and GSR), we also found a significant difference between the two groups in the “Extended Body questionnaire” ($p = 0.0043$; Congruent “C” group: $\mu = 0.25\%$, $std = 1.4$; Incongruent “I” group: $\mu = -1.28$, $std = 1.27$) suggesting that participants did not only feel ownership of the external object but also experienced a bias in the perception of their body boundaries (Figure 4.3 C Lower panel).

To conclude our analysis, we ran a Spearman Rank correlation to establish whether there was a relationship between the proposed ownership measures. We found a significant correlation between the scores on the Expanded Boundaries questionnaire and Agency ($p < 0.05$) as well as self-reported Ownership ($p < 0.001$). Finally, we report a significant relationship between Agency scores and the increase in the GSR signal ($p < 0.05$).

4.4 Discussion

Traditional approaches to body ownership are inspired by the Rubber Hand Illusion (RHI) [Botvinick and Cohen, 1998] in which the experience of ownership over a fake hand is induced by simultaneous stroking of the real hand occluded to vision and the rubber hand placed in a congruent position in front of the subject. A body of research which uses the RHI paradigm has argued that in the context when a participant is passively receiving sensory cues, body ownership depends on the bottom-up processes of multistory integration and top-down comparison between the predicted and the actual sensory stimuli based on prior knowledge [Botvinick and Cohen, 1998, Tsakiris and Haggard, 2005, Blanke, 2012, Apps and Tsakiris, 2014, Kalckert et al., 2019]. Thus, one might integrate a rubber hand into their body model, and experience ownership of it provided that it is physically, anatomically, posturally, and spatially congruent with their real hand [Tsakiris and Haggard, 2005, Costantini and Haggard, 2007, Lloyd, 2007, Haans et al., 2008, Makin et al., 2008, Ferri and Costantini, 2016]. Contrarily, in the context of action, that is, when subjects are voluntarily generating sensory feedback by, for instance, moving their arms, these conditions do not need to be satisfied. In fact, here, the effector neither needs to be connected to the body nor does it have to be physically plausible [Banakou et al., 2013, Peck et al., 2013, Romano et al., 2015, Van Dam and Stephens, 2018, Ma and Hommel, 2015a, Short and Ward, 2009, Maravita and Iriki, 2004]. It can be, for example, a two-dimensional digital balloon [Ma and Hommel, 2015b]. The only condition which needs to be met is that the behavior of this balloon follows the intentions of the agent and is fully predictable [Short and Ward, 2009, Ma and Hommel, 2015b].

This proposal is in line with the hypothesis that the experience of ownership is computed by the forward model, or corollary discharge, which guides motor behavior by estimating and correcting for the discrepancies between the predicted sensory consequences of movements and the actual afferent signals, the so-called sensory prediction errors [Miall and Wolpert, 1996, Wolpert and Kawato, 1998, Crapse and Sommer, 2008, Grechuta et al., 2019d]. According to this view, one could feel ownership of any effector, it being a hand, a tool, or a disconnected balloon as long as the sensory prediction errors are low. Crucially, forward models integrate across all sensory information which is relevant to the task and the goal at hand independently of whether it occurs within or outside of the peripersonal space. Until now, however, a vast majority of the studies support the notion that ownership requires the effector to be located in the peripersonal space [Brozzoli et al., 2011b, Maravita and Iriki, 2004] despite the evidence that, in the context of action, it is likely to result from the congruency of the forward model. The goal of this study was to address this question explicitly and to determine whether peripersonal space is a necessary condition for the experience of ownership over a non-corporeal object disconnected from the body, whose behavior can be controlled (via body movements) and therefore fully predictable. To this end, we designed a virtual reality-based protocol, where participants were required to control a virtual *log* and direct it towards assigned targets (fountains) by performing a set of five movements. We compared the subjective and objective experience of ownership between two conditions where the mapping between arm movements and the log was either congruent and therefore entirely predictable, or random and therefore not predictable and impossible to learn.

As expected, we found that in the congruent condition, participants experienced strong agency. Interestingly, however, our results also yielded a significantly stronger experience of ownership of the log in the congruent compared to the incongruent condition on both subjective and objective measures. In particular, in the condition where participants could control the log and predict its behavior, their responses to the threatening event, when the log was stubbed by an axe, were similar to as if it was their own hand being stubbed. Such a reaction has been commonly used to quantify the experience of the rubber hand illusion [Armel and Ramachandran, 2003]. Furthermore, we reported statistical differences in the Extended Body questionnaire between the two conditions such that the values were higher in the congruent condition. This result suggests that the control of the log might have biased the perceived limits of the boundaries of the physical body. Although this scale was designed specifically for this experiment and has not been previously validated, probably due to the novelty of the paradigm, the subsequent correlations between the Expanded Boundaries questionnaire and self-reported Agency as well as Ownership further support the consistency of the obtained results within different dimensions of ownership quantification. Crucially, these results did not depend on the virtual environment as such as in both groups participants reported the feeling of presence.

Similar to other studies [Short and Ward, 2009, Ma and Hommel, 2015b, Tsakiris et al., 2005], the results of this work support the notion that, contrary to the classical RHI [Botvinick and Cohen, 1998], in the context of voluntary goal-oriented control (of an effector) sensory prediction errors weight ownership stronger than top-down influences including the prior knowledge which comprises details of the specifics of the body. It might be that in the absence of control the forward model is *shut down* and therefore top-down control plays the major role in establishing the ownership by computing, for example, the similarity between the real and the rubber hands. Future studies will investigate the contribution of top-down processes to the experience of ownership by, for instance, hampering the access to attentional resources by introducing cognitive load in standard RHI on the one hand and in the context of action on the other.

Crucially, our results support that in action-contexts it is not required for the effector to be located within the peripersonal space. However, it needs to be noted that the overall self-reported experience of ownership in our and other similar studies [Ma and Hommel, 2015b] is generally weaker than in the classical RHI paradigms [Botvinick and Cohen, 1998] — an issue to be further experimentally explored in the future studies. We speculate, however, that different sensory events impact the ownership of any controlled effector in different ways which are subject to the history of sensorimotor contingencies of an individual. For instance, those events which are more frequently experienced and normally occur within the peripersonal *action* space can be predicted better and therefore result in stronger experiences of ownership than those experienced less frequently [Friston, 2010, Apps and Tsakiris, 2014, Suzuki et al., 2013].

From such perspective, any type of virtual object could be incorporated into the representation of one’s body provided that its behavior is controllable and predictable, that is, it is in accordance with the intentions of the agent [Short and Ward, 2009, Ma and Hommel, 2015b]. It might, however, require a long-enough exposure. Indeed such interpretation establishes a tight link between the experience of ownership (i.e., the sense that a given object or event is part of one’s body) and agency (i.e., the sense of control over a particular object or event) which, by now has not been resolved, possibly due to the diversity of experimental methods compared (passive and active rubber hand illusion) [Tsakiris et al., 2006b, Haggard, 2017, Sato and Yasuda, 2005, Ma and Hommel, 2015b, Liepelt et al., 2017, David et al., 2016, Kalckert and Ehrsson, 2012].

Together, our results extend current views on the processes underlying the experience of body ownership in action-contexts and challenge standard models of body ownership, which view ownership as a passive sensory state. These finding might find applications in the field of rehabilitation of motor functions as well as disturbances of ownership and phantom pain.

Chapter 5

CHALLENGING THE BOUNDARIES OF THE PHYSICAL SELF: PURELY DISTAL CUES IN THE ENVIRONMENT IMPACT BODY OWNERSHIP

This chapter is based on:
Grechuta, K., De la Torre, J., Rubio Ballester, B., & Verschure, P. (2019) Challenging the boundaries of the physical self: purely distal cues in the environment impact body ownership *Royal Society Open Science.* (Under review)

The unique ability to identify one’s own body and experience it as one’s own is fundamental in goal-oriented behavior and survival. However, the mechanisms underlying the so-called body ownership are yet not fully understood. The plasticity of body ownership has been studied using two experimental methods or their variations. Specifically, the Rubber Hand

Illusion (RHI), where the tactile stimuli are externally generated, or the *moving* RHI which implies self-initiated movements. Grounded in these paradigms, evidence has demonstrated that body ownership is a product of bottom-up reception of reafferent and exafferent multisensory information and top-down comparison between the predicted and the actual sensory stimuli. Crucially, provided the design of the current paradigms, where one of the manipulated cues always involves the processing of a proximal modality sensing the body or its surface (e.g., touch), the contribution of sensory signals which pertain to the environment remain elusive. Here we propose that, as any robust percept, body ownership depends on the integration and prediction of all the sensory stimuli, and therefore it will depend on the consistency of purely distal sensory signals pertaining to the environment. To test our hypothesis, we create an embodied goal-oriented task and manipulate the predictability of the surrounding environment by changing the congruency of purely distal multisensory cues while preserving bodily and action-driven signals entirely predictable. Our results empirically reveal that the way we represent our body is contingent upon all the sensory stimuli including purely distal and action-independent signals which pertain to the environment.

5.1 Introduction

The sense of body ownership, which allows us to determine the boundaries between the own physical self and the external world, and therefore the source of a given sensation, is fundamental in adaptive goal-oriented behavior and survival [Clark, 1999, Wolpert and Flanagan, 2001, Van Dam and Stephens, 2018, Crapse and Sommer, 2008]. Indeed, during the last three decades, scientists have increasingly questioned both the behavioral and neural mechanisms driving the emergence and experience of body ownership as well as its flexibility [Botvinick and Cohen, 1998, Tsakiris, 2010, Ehrsson et al., 2004, Blanke, 2012, Apps and Tsakiris, 2014, Ferri et al., 2017]. Together, the results support the notion that the way we perceive our body strongly relies on an interplay between (1)bottom-up reception, combination, and integration of self-generated (reafferent) and externally-generated (exafferent) information from multiple sensory sources, and (2)top-down comparison between the expected and the actual sensory stimuli [Botvinick and Cohen, 1998, Tsakiris and Haggard, 2005, Blanke, 2012, Apps and Tsakiris, 2014]. At the empirical level, the principles underlying bodily representation (in healthy subjects) have been studied using bodily illusions [Costantini, 2014]. A well-established experimental paradigm is the Rubber Hand Illusion (RHI) which is used to induce ownership over a fake rubber hand by manipulating the congruency of externally-delivered tactile cues [Botvinick and Cohen, 1998]. Another standard method for inducing ownership, in this case to a computer-generated body(part), is the so-called *moving* Rubber Hand Illusion (mRHI). Here, the visual feedback of self-initiated arm or finger movements is either synchronized with the actual trajectory or not, thus biasing the feeling of ownership over the virtual body [Tsakiris et al., 2006b, Dummer et al., 2009, Sanchez-Vives et al., 2010, Kalckert and Ehrsson, 2012]. Crucially, both in the RHI and mRHI or their variations, one of the sensory signals manipulated to induce the experience of ownership always involves processing of a proximal modality such as touch or proprioception [Botvinick and Cohen, 1998, Sanchez-Vives et al., 2010]. Consequently, the current understanding of the mechanisms driving body ownership is constrained to the study of sensory cues which pertain to the body or the sensory consequences of its self-initiated movements within the peripersonal space (PPS) [Rizzolatti et al., 1981a, Rizzolatti et al., 1981b, Holmes and Spence, 2004, Makin et al., 2007, Makin et al., 2008, Brozzoli et al., 2011a, Brozzoli et al., 2011b]. However, the contribution of purely distal signals outside of the PPS which pertain exclusively to the environment is still not fully understood.

With seemingly no effort we generate unambiguous interpretations about the self and the environment and determine the boundaries between them [Crapse and Sommer, 2008]. To do so, our brain uses multiple sources of sensory information processed by different modalities (i.e., vision, touch, audition) [Ernst and Bülthoff, 2004]. As such, any robust percept including the sense of ownership requires merging of this information which continuously occurs within and outside of the body, in the environment [Ernst and Bülthoff, 2004]. For instance, we simultaneously receive and integrate inputs informing about the location and position of our limbs as well as those informing about the location and position of objects in the scene. Until now, however, the experimental evidence about the multisensory representation of the body and the necessary and sufficient conditions for the experience of its ownership is grounded exclusively in the study of proximal cues [Botvinick and Cohen, 1998, Tsakiris, 2010, Blanke, 2012, Kilteni and Ehrsson, 2017, Ferri et al., 2017]. For instance, the seminal experiment of Botvinick & Cohen [Botvinick and Cohen, 1998] and later many others (for review see [Tsakiris, 2010]) who used RHI as a method to manipulate body ownership, proposed that the self-attribution of the rubber hand arises reactively as a result of bottom-up processes of combination and integration of information from visual and tactile modalities. Hence, originally the illusion of owning the fake hand was interpreted as a passive perceptual state whose strength was correlated with temporal discrepancies between seen and felt sensory stimuli (both necessary and sufficient). In the light of recent findings, however, this traditional view on body ownership as resulting purely from perceptual correlations does not seem sufficient [Tsakiris, 2010]. In particular, it has been widely accepted that in the context of externally-generated sensory cues (e.g., tactile strokes as in the RHI), body ownership strictly requires physical, anatomical, postural and spatial congruency of the real (felt) and fake (viewed) hands [Tsakiris and Haggard, 2005, Costantini and Haggard, 2007, Lloyd, 2007, Haans et al., 2008, Makin et al., 2008, Ferri and Costantini, 2016]. These findings strongly suggest that the interpretation of the 'novel' sensory evidence and possible incorporation of the rubber hand into the representation of the body is constrained by top-down prior knowledge [Tsakiris and Haggard, 2005, Blanke, 2012, Apps and Tsakiris, 2014]. In particular, the perception of ownership seems to rely on an internal model of the body which relates the physical aspects of the perceived rubber hand to the inputs received through a history of sensorimotor interactions of the agent with the world [Tsakiris, 2010, Apps and Tsakiris, 2014]. Interestingly, this hypothesis is consistent with the general framework which proposes that perception is controlled by top-down processes allowing to create predictions about the forthcoming sensory events based on previous experience and generalized knowledge [Engel et al., 2001, Friston et al., 2006, Kersten et al., 2004]. As such, it is an active process in which all acquired sensory information is continuously compared against experience-driven internal models of the self and the environment [Ferri et al., 2017, MacKay and Crammond, 1987, Perrett et al., 1985, Carlsson et al., 2000, Brown and Brüne, 2012].

Grounded in the framework of active perception, Ferri and colleagues [Ferri et al., 2013, Ferri et al., 2017] studied whether body ownership can be modulated by a pure expectation of exafferent tactile feedback in the absence of actual physical touch of either the fake or the real body-parts. Interestingly, their experiment revealed that a mere expectation of an upcoming sensory event, predicted by an anticipatory response in multisensory parietal cortices, is indeed sufficient to induce the experience of ownership over the rubber hand, measured subjectively and objectively. However, the tactile stimulation is not necessary [Ferri et al., 2013, Ferri et al., 2017] (see also [Smit et al., 2018]). This result emphasizes the predictive processing in the emergence of the sense of body ownership and challenges the traditionally defined boundaries of an embodied self. Here we extend this framework and propose that, as any coherent percept, body ownership is a result of bottom-up integration and top-down prediction of all the sensory stimuli processed by proximal and distal modalities including those which pertain purely to the environment. We, therefore, hypothesize that body ownership will depend on the consistency of distal sensory signals which occur in the environment and are independent of self-initiated actions. To test our hypothesis, we create an embodied goal-oriented task using virtual reality and manipulate the predictability of the surrounding environment by changing its rules while preserving bodily and action-driven signals fully predictable. We predict that body ownership will be influenced in the condition where purely external sensory signals underlying the statistical structure of the environment are not predictable.

5.2 Materials and Methods

5.2.1 Participants

Twenty-four healthy naive students from Universitat Politécnica de Catalunya provided their consent and participated in the study: 12 males (Mean±SD 23,25±2,37 years-old) and 12 females (Mean±SD 22,16±2,12 years-old). The sample size was chosen based on previous studies [Shergill et al., 2003, Kilteni et al., 2018]. All the subjects were right-handed (handedness assessed using the Edinburgh Handedness Inventory [Oldfield, 1971]), reported normal hearing and normal or corrected-to-normal vision. Each participant was pseudorandomly assigned to one of the two experimental conditions, following a between-subjects design. This experimental design prevented habituation to the ownership measures and sensory manipulations. The groups were balanced with respect to age, gender and previous experience with virtual reality. We ensured that none of the participants had previously participated in a body-ownership study, and informed the subjects that they were free to withdraw from the experiment at any time.

5.2.2 Experimental Setup, Procedures, and Protocol

Experimental Setup. The present experimental setup consisted of a personal computer, head mounted display (HTC Vive, www.vive.com), motion detection input device (Kinect, Microsoft, Seattle, USA), and active noise control headphones (Beats Electronics LLC, California, USA) which served to ensure isolation from external sounds (Figure 5.1 B). Similar to previous experiments [Sanchez-Vives et al., 2010,Grechuta et al., 2017b], here we used the virtual reality method as a tool to investigate the modulation of body ownership [Sanchez-Vives and Slater, 2005]. The protocol was integrated within the Rehabilitation Gaming System (RGS) [Grechuta et al., 2019d], the Virtual Environment (VE) was designed using SketchUp (Trimble Inc., California, USA) (Figure 5.1 C), and the software was developed in Unity3D (Unity Technologies SF, Copenhagen, Denmark). During the experiment, subjects sat at a cut-out table with their arms resting (Figure 5.1 B). The movements of the arms were continuously tracked and mapped onto the avatar's arms enabling interaction with the VE.

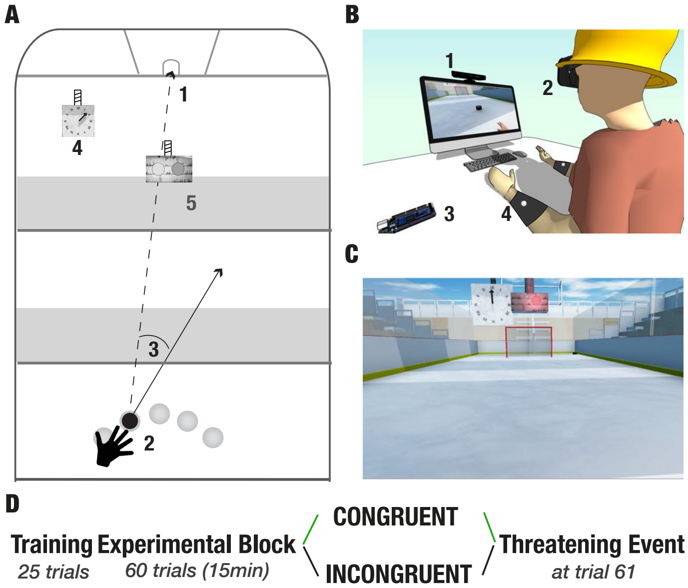


Figure 5.1: Experimental setup and protocol. **(A)Schematic bird view of the virtual environment.** 1- goal; 2- five starting positions; 3- an example of a directional error; 4- the Go Signal (GS); 5- a clock. **(B)Experimental setup.** 1- motion tracking device (Kinect); 2- Head Mounted Display (HMD, HTC-Vive Headset); 3- Arduino & E-Health; 4- Bracelets with reflecting markers for the tracking. **(C)Virtual scene.** The first person perspective view of the scene. **(D)Experimental protocol.** All participants go through the *Training Block* (25 trials). Then they are randomly split into two experimental conditions: congruent ("C", green arrow) and incongruent ("I", black arrow) and they undergo the *Experimental Block*. At the end of the experiment, all the participants experience the *Threatening Event*.

Virtual Air Hockey Task. The objective of each session was to complete a goal-oriented motor task that consisted of hitting a virtual puck into a goal, as accurately as possible (Figure 5.1 A). Prior to the experiment, subjects received instructions to place their hands in square-shaped Go Areas (GA; one for the left and one for the right hand) at the beginning of every trial. The trial started, and the puck appeared only when the system detected that both hands are in the GAs. Importantly, although both hands were mapped and rendered in the virtual scene, the task was to be completed using the right arm exclusively. To counteract repetitive movement-patterns and prevent habituation the puck was spawned pseudorandomly in one of the five Starting Positions (SP) distributed evenly within the right-hand workspace (Figure 5.1 A2). The puck did not bounce against the walls. Thus a trial consisted of one hit only which could end in either a success (the puck enters the goal) or a failure (the puck hits one of the walls). Both events were immediately indicated by auditory feedback in the form of semantically corresponding sound, positive or negative respectively. Pertaining to the task, this feedback was always congruent and fully predictable.

Protocol and Sensory Manipulations. The experiment consisted of three parts including (1)baseline Training Block (TB, 25 trials), (2)Experimental Block (EB, 60 trials), and (3)a Threatening Event (TE) (Figure 5.1 D). TB and TE were the same for all the participants. TB, in particular, allowed familiarization with the virtual environment as well as with the dynamics of the game, while TE served to record autonomous, physiological responses to an unexpected threat as an objective measure of body ownership [Armel and Ramachandran, 2003]. In the EB, subjects were randomly split into two groups: Congruent “C” (green), and Incongruent “I” (black) (Figure 5.1 D). To investigate whether sensory cues which pertain to the environment influence body ownership in the experimental condition (“I”), we manipulated the congruency, and therefore the predictability, of visual and auditory action-independent cues from the virtual scene.

The scene consisted of (1) the virtual arms, (2) an air hockey field, (3) the puck, (4) a goal, (5) bleachers for the audience, (6) the Go Signal (GS), and (7) a clock (Figure 5.1). The virtual solar time was indicated by the position of the sun in the sky (i.e., visual cue), while the virtual space (setting) was signaled by the background sound representative for a given place (i.e., auditory cue; e.g., the sound of the air-hockey field) (Figure 5.1 C). The default scene was set at midday (setting:time) on a hockey field (setting:location). Both in the Training Block (identical for both conditions) and the Experimental Block of the Congruent condition "C" all the scene components mentioned above, as well as the temporal and spatial settings, remained fixed such that their behavior was always fully predictable. Moreover, in all conditions, all the auditory and visual signals relevant to the body within the peripersonal space (the mapping between the real and the the virtual arms) and to the task (air hockey field, the puck, the goal, and the Go Signal, the trajectory of the puck, outcome of the action) were always congruent and fully predictable. Crucially, in the Experimental Block of the Incongruent condition ("I"), the default behavior of the scene components and the temporal and spatial settings randomly changed. In particular, we manipulated: (1) spatial orientation of the bleachers by rotating them on the z-axis, (2) spatial orientation of the clock by modulating the velocity and the direction of the arrows indicating the time, (3) virtual solar time by altering the position of the sun in the sky or changing its characteristics to nighttime, and (4) virtual space by altering the background sound (i.e., sounds representative for a concert, cinema). Importantly, to ensure that the sensory manipulations in "I" impact exclusively the perception of the environment and not the action, they were always introduced between the end of a trial (the puck enters the goal or hits one of the walls) and the beginning of the consecutive one. Specifically, they were triggered at a random time within a 2 seconds time window after the end of each trial. The incongruencies were introduced gradually, and they were pseudorandomly distributed such that the participants could not attribute action-driven causality to their emergence.

5.2.3 Measures

Self-reports. In virtual environments, the sense of presence refers to the subjective experience of 'being there', despite the physical distance. In particular, when a user does not perceive the influence of technology during a virtual reality-based experience [Witmer and Singer, 1998,Sanchez-Vives and Slater, 2005]. To ensure that the participants in both groups felt equally immersed within the proposed environment, we asked them to complete a presence questionnaire at the end of the experiment by assessing each of the items on a 9-point Likert Scale (see Table5.1 for the full list of items). Furthermore, to evaluate the subjective experience of body ownership and agency, we administered a 12 item questionnaire adapted from previous studies [Longo et al., 2008, Kalckert and Ehrsson, 2012]. There were six questions per domain, three of which served as controls. Participants answered each statement on a 7-point Likert Scale ranging from ‘-3’: being in strong disagreement to ‘3’: being in strong agreement. To counteract order effects, the sequence of questions was randomized.

Galvanic Skin Response (GSR). GSR is a physiological measure of the autonomous nervous system, which increases as a reaction to an arousing stimulus. Similar to other studies [Armel and Ramachandran, 2003], here, we used GSR as a proxy for the ownership illusion. In particular, at the end of the experimental block in each condition, we measured GSR responses to an unexpected threat (i.e., a knife falling to stab the right virtual hand). To prevent movement-driven muscular artifacts, we recorded GSR from the left hand which did not move during the experiment. The signal was recorded using Arduino e-Health board at a sampling rate of 33Hz from two flat reversible silver/silver chloride (Ag-AgCl) electrodes which were attached to the middle and index fingers, respectively. We stored the GSR during the whole experiment interval for each participant. The data were preprocessed to extract phasic components from tonic activity based on Continuous Decomposition Analysis (CDA) [Benedek and Kaernbach, 2010b] as implemented in the Ledalab software (Leipzig, Germany). For the analysis, we computed an increase in the number of galvanic skin responses (nGSR) above 0,01mS 1 second following the threatening event and compared it between the two conditions.

Hand Withdrawal (HW). We collected kinematic movement data from the Kinect for each participant throughout the experiment. All the data from the system was recorded at 33Hz. To quantify the execution of instinctive defensive movements such as hand withdrawal in response to the unexpected threatening event (i.e., the virtual knife stabbing the virtual hand) [González-Franco et al., 2014], we computed the velocity of displacement of the right virtual hand as a difference in the cumulative sum of the X (forward and backward) and Y (up and down) positions at every time step. The results were compared between the two conditions. Due to possibly stronger assimilation of the virtual hand to the representation of the body, we expected higher velocity of movement in “C” than in “I”.

Performance measures. To evaluate performance, we measured scores, angular errors as well as reaction and response times, all stored by the system. Scores were calculated as the percentage of successful trials, namely, the times when the puck entered the gate (Figure 5.1 A1). An angular error was computed as the difference between the actual direction vector and a straight line between the starting position of the puck and the middle of the goal (desired trajectory, Figure 5.1 A3). Reaction times were the time intervals between the apparition of the puck and the moment of ‘leaving’ the starting position to hit it, while the response times were the time intervals between the apparition of the puck and the moment of its collision with the hand.

The statistical analysis followed nonparametric methods. Hence, we used the Mann-Whitney U test for between groups analyses and the Wilcoxon signed-rank test for within groups comparisons. The data were analyzed using Python3.6.4 (http://www.python.org) and Matlab (Mathworks, Natick, USA).

5.3 Results

To determine whether body ownership depends on the consistency of purely external cues which pertain to the environment, we developed an embodied goal-oriented task and manipulated the congruency, and therefore the predictability of action-independent sensory cues outside of the peripersonal space while preserving bodily and action-driven cues fully predictable. For each session and each participant in both conditions, we used Galvanic Skin Response (GSR), Hand Withdrawal (HW), and self-reports to quantify body ownership objectively, behaviorally and subjectively. Furthermore, we stored performance scores, angular errors as well as response and reaction times as measures of performance. Finally, to ensure that the participants were immersed in the proposed virtual environment and perceived the sense of agency over the virtual limbs, we administered *presence* (adapted from [Witmer and Singer, 1998]) and *agency* questionnaires (adapted from [Longo et al., 2008]).

Presence and agency. First, we assessed the perceived experience of presence and agency. The analysis revealed that in both conditions participants felt present in the proposed virtual environment (“C”: $\mu = 1.6, std = 1.56$ and “I”: $\mu = 1.51, std = 1.8$) (Figure 5.2). Crucially, we found no differences between the groups in the self-reported scores ($p = 0.47$). Table5.1 presents individual questionnaire items as well as the results of between-group analyses for the associated questions. None of the comparisons yielded a statistically significant difference. We further report no difference in the experienced sense of agency between “C” ($\mu = 1.19, std = 1.24$) and “I” ($\mu = 1.3, std = 1.3$) ($p = 0.08$) (Figure 5.2), and the groups did not differ in the agency control questions ($p = 0.1$).

Performance. Our results revealed that the normalized performance-scores (i.e., the proportion of successful trials) did not differ between the Congruent ($\mu = 0.6, std = 0.17$) and the Incongruent ($\mu = 0.59, std = 0.18$) conditions ($p = 0.5$). We further analyzed the angular error as a proxy to performance, stored throughout the experiment. We report that while both groups significantly improved during the Training Block (TB), in the Experimental Block (EB) the errors stabilized (Figure 5.3). Specifically, there was a statistically significant difference between the early and late trials in both “C” (early trials: $\mu = 12.63, std = 5.25$ and late trials: $\mu = 6.11, std = 3.76; p = 0.002$) and “I” (early trials: $\mu = 14.03, std = 6.9$ and late trials: $\mu = 7.53, std = 2.48; p = 0.004$) in the training block. We found,

Table 5.1: All items from presence questionnaire. P-values indicate the results of a between-group comparison all the items using the Mann-Whitney U test.

question	**p-value**
Were you able to control events?	0.179
How responsive was the environment to actions you initiated?	0.257
Naturality of the interaction?	0.232
How much did visual aspects involve you?	0.476
How natural was the movement mechanism?	0.34
How compelling was the sense of objects moving?	0.21
How consistent were the experiences with real world ones?	0.384
Were you able to anticipate consequences of your actions?	0.22
Were you able to survey the environment using vision?	0.39
How compelling was the sense of moving around?	0.36
How closely were you able to examine objects?	0.22
Were you able to examine objects from multiple viewpoints?	0.2
How involved were you in the virtual reality experience?	0.18
How much delay did you experience?	0.476
How quickly did you adjust to the virtual reality experience?	0.4
How proficient in interacting did you feel at the end?	0.439
How much did visual aspects distract from the task?	0.38
How much did the control devices interfere with performance?	0.373
How well could you concentrate on the assigned tasks?	0.164
How much did auditory aspects involve you?	0.064
How well could you identify sounds?	0.474
How well could you localize sounds?	0.21

however, no within-group differences for the early and late trials in the experimental block: "C" (early trials: $\mu = 8.05, std = 3.81$ and late trials: $\mu = 7.14, std = 4.87; p = 0.53$) and "I" (early trials: $\mu = 7.1, std = 3.85$ and late trials: $\mu = 7.26, std = 3.94; p = 0.48$). Furthermore, the Mann-Whitney U test yielded no differences in performance (angular error) between "C" and "I" in neither the TB ($p = 0.15$) nor the EB (N=10, $p = 0.15$) (Figure 5.3) demonstrating that, overall, the conditions did not differ with respect to performance. Finally, our analysis showed no statistically

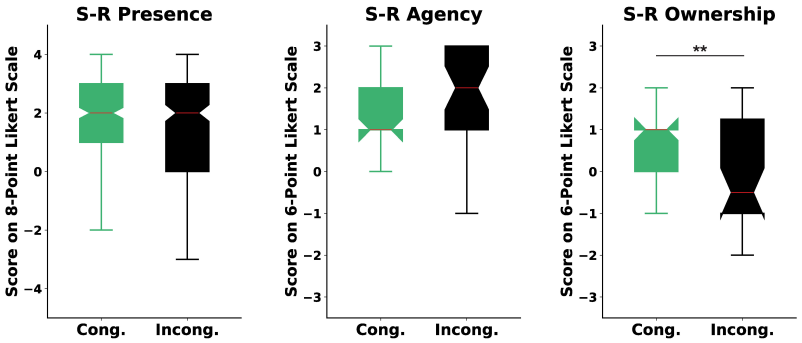


Figure 5.2: Self Reports. **Left.** Self-reported experience of presence. Y-axis: Responses on the 9-point Likert scale ranging from ‘-4’ (strongly disagree) to ‘4’ (strongly agree). Scores above ‘0’ indicate a feeling of presence. **Middle.** Self-reported experience of agency. Y-axis: Responses on the 6-point Likert scale ranging from ‘-3’ (strongly disagree) to ‘3’ (strongly agree). Scores above ‘0’ indicate a feeling of agency. **Right.** Self-reported experience of ownership. Y-axis: Responses on the 6-point Likert scale ranging from ‘-3’ (strongly disagree) to ‘3’ (strongly agree). Scores above ‘0’ indicate a feeling of body ownership.

significant differences in either response (“C”: $\mu = 2.35$, $std = 0.78$ and “I”: $\mu = 2.43$, $std = 0.67$; $p = 0.1$) or reaction times (“C”: $\mu = 1.01$, $std = 1.41$ and “I”: $\mu = 1.63$, $std = 1.14$; $p = 0.3$) between the two conditions during the experimental block. Hence both groups took the same time to initiate the movement and hit the puck, further suggesting that the proposed manipulation of action-independent sensory signals in “I” did not alter or bias performance.

Body Ownership. The analysis revealed a statistical difference in the self-reported experience of body ownership between the two conditions ($p = 0.04$) such that the scores were significantly higher in “C” ($\mu = 0.6, std = 0.9$) than in “I” ($\mu = 0.05, std = 1.45$) (Figure 5.2). Importantly, we found no differences in the control questions between the groups ($p = 0.1$). To further explore the effects of the congruency of purely external cues on body ownership, we computed baseline-subtracted post-threatening GSR responses for every individual in both groups (Figure

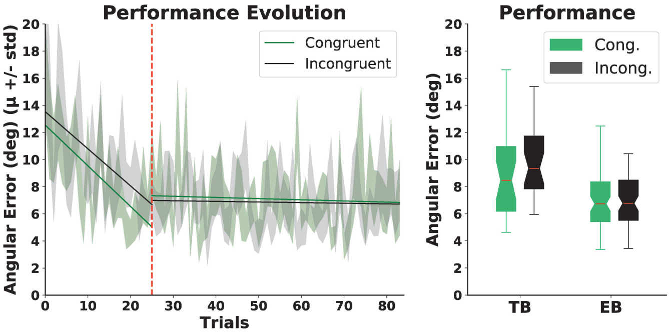


Figure 5.3: Motor performance. **Left.** The Evolution of Angular Errors. The dashed red line indicates the end of the Training Block (TB) and the beginning of the Experimental Block (EB). The solid lines represent linear regression models for the angular errors in each condition in the training block and experimental blocks, before or after the dashed red line respectively. **Right** Total Angular Errors. Boxplots represent angular errors for the two conditions in the training and experimental blocks respectively. No differences were found between the groups.

5.4). As expected and in line with the literature, the GSR signal increased in both groups. Specifically, the threatening event triggered a significant increase in the number of galvanic skin responses in “C” (pre-TE: $\mu = 1, std = 0.7$, post-TE: $\mu = 4.5, std = 2.14, p = 0.002$) and in “I” (pre-TE: $\mu = .83, std = 0.68$, post-TE: $\mu = 2.33, std = 1.64, p = 0.03$). Crucially, however, we found a statistically significant difference between the groups in the numbers of activations ($p = 0.009$) such that the number was significantly higher in “C” ($\mu = 4.5, std = 2.14$) than in “I” ($\mu = 2.33, std = 1.64$) (Figure 5.4).

Finally, we observed that in the congruent condition participants exhibited faster velocity of the right virtual hand displacement post threatening event (Hand Withdrawal, Figure 5.5). In particular, the statistical analysis revealed that the difference between “I” and “C” in the cumulative sum of the X and Y position over time reached statistical significance at second 4 post threatening event (Figure 5.5).

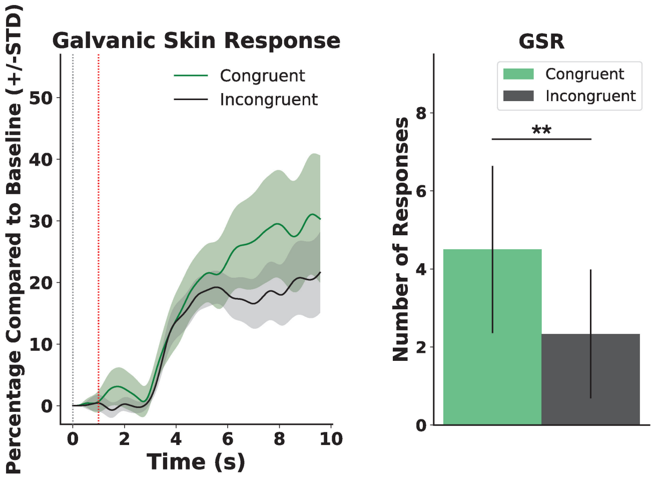


Figure 5.4: GSR results. **Left.** Evolution of the GSR. Responses per condition for all the participants averaged in a time window of 10 seconds. The post Threatening Event (TE) signals were normalized for each participant by subtracting the mean signal from 10 seconds prior to the stimulus onset. The dashed gray line indicates the time when the knife appeared ($time = 0$) while the dashed red line shows the time when the knife stabbed the hand ($time = 1$). The sampling rate for the signal was 33Hz. The data were run through a low-pass filter with a cut-off frequency of 3Hz. **Right.** Number of Galvanic Skin Responses. The plot represents the difference between the groups in the number of galvanic skin responses (nGSR) post stabbing event.

5.4 Discussion

The unique ability to recognize one's own body, experience it as its own, and localize it in space lies in the continuous processing of reafferent and exafferent multisensory information arising from sensorimotor interactions of an agent within the environment [Kawato, 1999, Clark, 1999, Wolpert and Flanagan, 2001, Crapse and Sommer, 2008]. Vast evidence has now demonstrated that this processing comprises bottom-up reception and top-down prediction of sensory stimuli which pertain to the body and occur within the peripersonal space [Botvinick and Cohen, 1998, Ferri et al., 2013, Sanchez-Vives et al., 2010, Maravita and Iriki, 2004, Serino et al.,

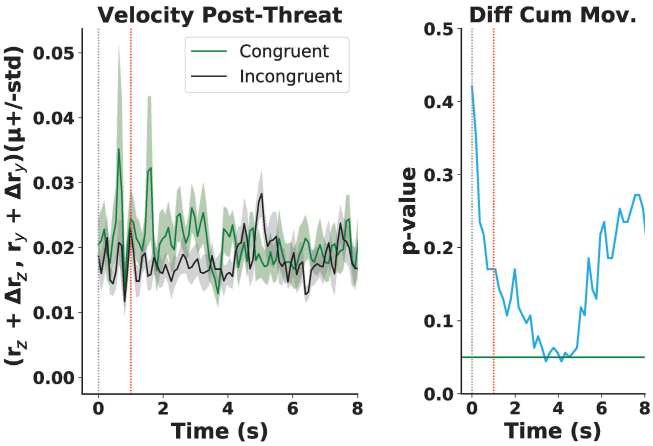


Figure 5.5: Hand Withdrawal results. **Left.** Changes in Movement Velocity Post-Threatening Event. The graph represents the evolution of changes in X and Y positions of the right virtual hand following the threatening event for each group over time. The data on the X-axis was recorded at 33Hz, and it is presented in an overlapping window of 66ms. **Right.** Changes in Velocity Between the groups. The blue line represents $p-values$ obtained from Mann-Whitney U tests performed on the cumulative sum of the movement velocity. The values corresponding to each time step on the X-axis are shown in overlapping windows of 150-ms. In B, the green horizontal line indicates 0.05 significance threshold. In both graphs, the dashed gray line indicates the time when the knife appeared ($time = 0$) while the dashed red line shows the time when the knife stabbed the hand ($time = 1$).

2015, D’Angelo et al., 2018]. The present study extends prior findings by showing that the plasticity of body ownership also depends on the consistency of body and action-independent sensory cues which pertain to the environment and are processed by purely distal modalities (i.e., visual and auditory). Thus, for the first time, we empirically reveal that the way we represent our body is contingent upon all the sensory stimuli including signals occurring outside the peripersonal space. We interpret our results from the perspective of active perception and propose that, similar to any robust percept, body ownership depends on the consistency of the internal models of not only the body or the consequences of its actions but also the model of the surrounding environment [Engel et al., 2001, Miall and

Wolpert, 1996, Crapse and Sommer, 2008, Ernst and Bülthoff, 2004].

A large body of evidence demonstrates that the experience of ownership emerges actively through dynamic comparisons between integrated and predicted multisensory signals [Tsakiris, 2010, Ferri et al., 2017, Grechuta et al., 2019d, Apps and Tsakiris, 2014, Seth, 2013, Suzuki et al., 2013]. The influence of top-down processing [Suzuki et al., 2013] on the sense of ownership is supported in the contexts of classical Rubber Hand Illusion (RHI) where the stimuli are externally generated [Botvinick and Cohen, 1998]. Here, the self-attribution of the rubber hand arises actively as a consequence of the minimization of prediction errors resulting from multisensory conflicts during the synchronous stroking of the real and fake body-parts (i.e., visuotactile) [Apps and Tsakiris, 2014]. Similar, a failure to experience ownership over a noncorporeal object [Tsakiris et al., 2010] or a rubber hand located in an implausible position [Costantini and Haggard, 2007] or of a different color [Farmer et al., 2012] can also be interpreted as a consequence of predictive matching of the sensory inflow and the experience-driven internal models of the body [Apps and Tsakiris, 2014]. Moreover, it has been recently demonstrated that ownership can be induced by a pure expectation of correlated sensory input [Ferri et al., 2013, Ferri et al., 2017, Smit et al., 2018].

The top-down processing in the emergence of ownership is further supported in the context of self-generated cues such as in the *moving* Rubber Hand Illusion (mRHI) [Dummer et al., 2009, Sanchez-Vives and Slater, 2005]. Here, it has been proposed that the location of different body-parts is estimated by the Central Nervous System (CNS) via a Forward Model (FM) or a Corollary Discharge (CD) which generates predictions about the sensory consequences of movements and compares them with the corresponding sensory feedback [Wolpert et al., 1995, Miall and Wolpert, 1996, Sommer and Wurtz, 2008]. Those predictions are carried out by the so-called efference copy, and they employ all the sensory signals relevant to the body and the goal of the task (task-relevant) [Griisser, 1995, Sperry, 1950, Wolpert et al., 2011]. The Sensory Prediction Errors (SPE) from multiple sensory sources, which reflect the discrepancies between the expected and the actual sensory stimuli, inform the brain about the current state of the environment and the body, shaping the experience of ownership. For instance when the visual feedback of the virtual hand does not match the expected one (the SPEs pertaining to the body are high) [Dummer et al., 2009, Sanchez-Vives et al., 2010] or the prediction about the auditory feedback of a puck hitting a wall is violated (the SPEs relevant to the task are high) [Grechuta et al., 2019d], the ownership of the virtual body decreases.

The evidence discussed above suggests that body ownership is compromised when the actual sensory signals violate the expected cues independently of whether they are externally (RHI) or self-generated (mRHI). If body ownership depends on the matching between the predicted and the actual sensory stimuli, can it, in a similar way, be affected by prediction errors about the sensory signals pertaining to the environment? To answer this question, we designed a virtual reality-based paradigm where participants were to complete a motor task (air hockey) and manipulated the predictability of the purely external cues by randomly changing the rules of the environment. Thus, similar to the prediction errors which result from visuotactile matching and affect the internal (generative) model of the body [Botvinick and Cohen, 1998], or those which result from visuomotor [Sanchez-Vives et al., 2010] or visuoauditory [Grechuta et al., 2019d] matching and affect the internal (forward) model, here we experimentally biased prediction errors which result from visuoauditory matching and affect the internal (generative) model of the environment [Clark, 2013, Friston, 2010]. We expected that if body ownership depends on the congruency of all the sensory stimuli, it will be impacted in the experimental phase of the incongruent condition where the expectations about the model of the environment acquired during the training phase are violated. Our findings establish that incongruencies in action-independent and task-irrelevant sensory cues, which inform about the statistical structure of the environment and are processed by purely distal modalities, modulate the experience of body ownership. In particular, we found that the congruent (as compared to incongruent) environment led to an enhanced experience of ownership over the virtual hand, as measured subjectively by a questionnaire (Figure 5.2), behaviorally using the hand withdrawal (Figure 5.5), and objectively through the galvanic skin responses (Figure 5.4). Crucially, however, there were no effects regarding the experience of presence (Table 5.1 and Figure 5.2) supporting that, despite the introduced manipulations, the environment was overall immersive [Sanchez-Vives and Slater, 2005].

We propose that the violation of expectations in the context of the proposed paradigm can be understood as a sudden increase of uncertainty in the internal model of the environment. According to biologically-constrained models of the neocortex, the link between these two components could be mediated by neuromodulators which signal uncertainty such as norepinephrine or acetylcholine [Angela and Dayan, 2005]. Consequently, a significant enough violation of expectations influenced by the sensory manipulations in the incongruent condition would have a global effect on modulating uncertainty in a range of brain areas including those which underlie the multisensory representation of the body (e.g., the right insula, posterior parietal and ventral premotor cortices) [Craig and Craig, 2009, Tsakiris et al., 2006a, Gentile et al., 2013]. From the functional perspective, such sensory prediction errors would have the same impact on all predictive models inducing uncertainty in the model of the environment (i.e., a generative model) and the model of the body. In other words, the temporal modulation of global uncertainty would inevitably change the overall confidence in the internal model of the body as supported by our results (Figure 5.2, 5.4, 5.5). The present outcomes are further consistent with the Bayesian causal inference and could be interpreted in terms of likelihood according to which if the environment is likely *I* am likely too, and vice-versa [Körding and Wolpert, 2004, Doya et al., 2007, Apps and Tsakiris, 2014, Samad et al., 2015].

An interesting question which one could raise, however, is how persistent is this effect? Is it transient? It has been demonstrated that one way to minimize prediction errors is to update the current model, in our case the model of the environment, to accommodate the unexpected sensory signals [Körding and Wolpert, 2004, Friston et al., 2010]. This would suggest that prolonged exposure to random errors would lead to a subsequent reduction of uncertainty in the model of the environment reducing the neuromodulatory response, which in turn would reduce the uncertainty in the model of the body. The consequent increase in the reliability of the predictive models of the environment and the body would immediately result in the reestablishment of body ownership. In particular, we expect that after more prolonged exposure to the incongruent stimuli in “I”, the experience of ownership as measured by questionnaires, hand withdrawal, and GSR would return to baseline such that there would be no differences in the perceived ownership between the two conditions. Further work shall systematically address this question by running additional trials to assess the temporal evolution of body ownership in the context of incongruent environment.

What about performance? Our results demonstrate that the sensory manipulations in the incongruent condition did not affect either the self-reported experience of agency (Figure 5.2) or the performance measured through scores, angular errors (Figure 5.3) or response times. On the one hand, we designed the paradigm such that all the bodily (i.e., within the peripersonal space), action-driven (reafferent) and task-relevant stimuli were always congruent and therefore fully predictable. Specifically, the 1:1 mapping between the real and the virtual hands ensured that the visual feedback of the movements in virtual reality fully reflected the movements of the real hands resulting in visuomotor congruency. Similar to the peripersonal signals, the consequences of actions in the extrapersonal space signaled by auditory and visual feedback reflected real-world physics and were fully predictable. That is, the sound of the puck temporarily and spatially corresponded to the location of its collision with the environment (i.e., the walls or the goal). On the other hand, we also experimentally controlled for the occurrence of the sensory manipulations to ensure that they are action-independent. Specifically, they were always introduced randomly between the end of a trial and the beginning of the next one. Finally, neither did the manipulated signals in “I” inform about the outcomes of the task (i.e., knowledge of performance) nor did they affect motor performance, which made them task-irrelevant. Indeed, the goal of every experimental session was to complete the motor task as accurately as possible by hitting the puck into the goal. We thus speculate that the congruency (and therefore the predictability) of all the sensory information relevant to the effector and the target in both conditions resulted in an unbiased performance and reinforced the experience of agency, that is the experience of controlling one’s actions, and, through them, events in the outside world [Wolpert et al., 2011, Haggard, 2017].

In conclusion, our results support the notion that the plasticity of body ownership depends on an active interplay between the experience-driven top-down predictions and bottom-up prediction errors driven by purely external and action-independent cues which pertain to the environment. Hence, these findings extend current accounts by demonstrating that the sensory evidence necessary for constructing ownership goes beyond the *body* and the peripersonal space [Makin et al., 2008]. In line with the motor control and perception studies, our data support a functional coupling between the predictive (generative) models of the body and environment [Miall and Wolpert, 1996]. Future work should include a systematic study of the weighting of specific exafferent and reafferent unimodal and multisensory information in modulating the experience of body ownership under different tasks as well as their neural underpinnings. At this point, however, we expect that the current findings will allow for the advancement of our understanding of the principles underlying the emergence and experience of body ownership, which we propose can be understood in a framework of active inference of all the signals within and outside of the peripersonal space. We believe that the reported results can also contribute to the development of robust computer-based paradigms for the treatment of neurological disorders of perceptual and motor functions such as somatoparaphrenia [Banks et al., 1989, Cameirão et al., 2010].

Part II

Chapter 6

INTENSIVE LANGUAGE-ACTION THERAPY IN VIRTUAL REALITY FOR A REHABILITATION GAMING SYSTEM

This chapter is based on:
Grechuta, K., Ballester, B. R., Duff, A., Oller, E. D., Pulvermuller, F., & Verschure, P. F. (2016). Intensive language-action therapy in virtual reality for a rehabilitation gaming system. *Journal of Pain Management, 9*(3), 243.

One third of stroke patients suffer from language disorders. Recently, Intensive Language Action Therapy (ILAT) emerged as a novel paradigm for aphasia rehabilitation. In the present study, we designed and developed a virtual reality (VR) based language rehabilitation tool by integrating ILAT’s object request Language Action Game (LAG) in a Rehabilitation

Gaming System (RGS), a novel paradigm for the rehabilitation of motor deficits after lesions to the central nervous system. RGS is an environment that provides multimodal, task specific training in virtual reality scenarios. Its special design consists of a motion detection system that monitors users' movements, which allows for an active interaction, as well as continuous evaluation of the performance. We address the question of whether aphasia rehabilitation designed within the VR environment of RGS can be effective. We report the results of a double-case pilot study where one acute and one chronic aphasic patient followed five RGSILAT therapy sessions. Before and after the treatment, we evaluated their language skills using the Communication Activity Log (CAL) and Western Aphasia Battery (WAB) scales. Results show that the patients learnt how to interact within the VR system. The CAL performance suggests that both patients and their therapist perceived improvements in communication skills after the intervention. Additionally, both approval and acceptance of the system were high. Based on this initial outcome we will provide the present RGS-ILAT with further advancements and evaluate the system with higher number of patients.

6.1 Introduction

Stroke is a neurological disease which causes the most common disabling neurological damages [Carter et al., 2012]. 35-40% of stroke patients suffer serious language deficits, such as aphasias, which are often accompanied by anxiety, depression or social withdrawal [Elman et al., 2000]. Traditional aphasia therapies focus mostly on repeating words, where the complexity of the practiced language gradually changes from less to more frequent. These methods usually do not put emphasis on the importance of an intense practice of language adapted to the personal needs of each patient, within a meaningful context. Alternative treatment and rehabilitation methods are therefore required in order to achieve successful recovery.

Recently, the relation between language, cognition and its neural substrate has shed light onto the composite structure of the language processing and production systems as well as the effective rehabilitation of language deficits caused by stroke [Özyürek et al., 2007, Pulvermüller, 2005, Lakoff and Johnson, 2008]. The brain comprises a set of interconnected neural circuits where linguistic, or any other, motor, perceptual or attentional abilities cannot be separated into discrete modules [Carter et al., 2012]. Therefore, for a therapy to be effective, in the brain there must be an interaction between linguistic neural system, motor and sensory circuits, memory, planning and monitoring [Kurland et al., 2012]. It has been shown that both words and complex sentences, which are semantically related to actions that involve different parts of the body, activate the sensorimotor cortex [Pulvermüller, 2005, Berthier and Pulvermüller, 2011]. This observation has led to the hypothesis that sensorimotor circuits provide the cortical basis for language [Pulvermüller, 2005]. Accordingly, language processing, both comprehension and production, is physically linked to the action systems. This is consistent with the general view on the tight coupling of sensing and action in the brain [Verschure et al., 2003]. Being embedded within the sensorimotor system, language processing is coupled to one’s bodily experience, which suggests a novel route for the rehabilitation of language deficits. Indeed, it has been reported that a specific action oriented language training can result in considerable improvements in both language performance and its underlying cerebral activity related to language, in both Wernicke’s and Broca’s aphasia patients [Pulvermüller, 2005]. The research on the reorganization of language related brain areas, which follows rehabilitation, suggests that neural plasticity and reorganization can even result in shifts in language lateralization [Neville et al., 1998]. These findings have changed the approach towards the language rehabilitation reinforcing the stimulation of multiple brain regions creating conditions for recovery [Carter et al., 2012]. The range and types of language rehabilitation techniques have been further amplified by using a range of technologies including virtual reality tools, which have shown to be successful in treating deficits resulting from stroke [Abad et al., 2013, Cameirão et al., 2007, Cameirao et al., 2012]. In particular, we have shown previously that an approach, which combines mirroring through VR with specific brain-theory based training protocols, or the Rehabilitation

Gaming System (RGS), can be highly effective in the rehabilitation of the upper extremities in acute and chronic stroke patients [Cameirão et al., 2007,Cameirão et al., 2010,Cameirao et al., 2012]. Here, we further extend this RGS approach by augmenting it with a VR based version of ILAT. In particular we investigate the question whether RGS-ILAT is effective in treating stroke induced Broca’s aphasics [Difrancesco et al., 2012].

6.1.1 ILAT and Broca’s Aphasia

ILAT is a Speech and Language Therapy (SLT) approach that aims at reinforcing the activation of both linguistic and its underlying motor circuits in a systematic and structured way by means of intensive practice and contextualized game scenarios [Pulvermüller, 2005]. The therapy focuses on treating Broca’s aphasia that results from the lesion to the left frontal cortex [Boo and Rose, 2011]. The syndrome is characterized by disorders in the syntax of language production including motor disorders and agrammatism. Individuals who suffer this type of aphasia are typically not fluent when speaking and often cannot combine words into meaningful sentences [Marshall, 2008]. Patients who suffer from Broca’s aphasia therefore may benefit from rehabilitation methods that focus on the reinforcement of full sentence production as well as general fluency. Within this context, Pulvermüller et al. emphasize three main premises of ILAT [Difrancesco et al., 2012]. The first one is the intensive training (e.g. 3h/week for 2 weeks). Secondly, ILAT exploits the behavioral relevance of the therapeutic context, namely, the embodiment of speech within a communicative, natural, action context. Finally, the authors suggest the use of behavioral techniques such as shaping, modeling and positive reinforcement. Indeed, recent studies show that even patients with severe Broca’s aphasia and/or Apraxia of Speech (AOS) can improve when undergoing ILAT [Kurland et al., 2012]. In the present study we propose a new rehabilitation scenario that combines ILAT and RGS. We believe that the original ILAT may benefit from its VR implementation, which allows for the implementation of multimodal feedback, and provides wider accessibility.

6.1.2 ILAT Scenario

There are 3-4 players who take part in the original ILAT session. One of the players is a Speech and Language Therapist (SLT), whose role is to actively monitor the patients, keep track of utterances, model speech and adjust the velocity of the game. The rest of the participants are patients with post-stroke Broca’s aphasia. The Original ILAT consists of two types of Language Action Games (LAGs): the object request LAG (see Figure 6.1) and the action-planning LAG [Difrancesco et al., 2012]. The object request LAG begins when all the participants are given identical sets of cards (from 6 to 12 each). The player who starts the game (player A) selects one card and holds it in his/her hand, so that the other participants cannot see its content. Next, s/he verbally requests the same card from another player (player B). The possible moves that can follow depend on whether the player B owns the requested card. Player B can therefore either follow the request and pass the corresponding card, or reject the request. Further clarification attempts can occur in case of misunderstandings between the players. The goal of the object request LAG is for the player to be the first with no cards left on the table. This can be achieved by either passing or receiving the matching card/s. In the present project we have implemented the object request LAG protocol in RGS by rigorously following its language-action structure.

6.1.3 RGS

RGS is a novel paradigm for the rehabilitation of motor deficits after lesions to the central nervous system [Cameirão et al., 2010]. It is a gaming environment that provides a multimodal, task specific training in virtual reality scenarios 6.2. Its special design consists of an intelligent motion detection system that monitors the users’ movements. This allows for an active interaction as well as continuous evaluation of the affected limbs. The original purpose of the system is to provide a novel rehabilitation tool to treat motor deficits of upper limbs in post stroke patients. RGS deploys a number of scenarios which can be adjusted to the specific needs of the users. That allows for a continuous interaction with the Virtual

Environment (VE). The computer-generated world is viewed form the first person's perspective, and all the events that happen within the VR are under realtime user control. RGS has proved to be successful in the number of clinical trials [Cameirão et al., 2007, Cameirão et al., 2010, Cameirao et al., 2012]. We see RGS as an example of the novel field of science-based medicine where interventions are based on causal theories of brain and behavior. Its tracking system, individualized training and reinforced visual feedback [Cameirão et al., 2010] allow for the integration of ILAT.

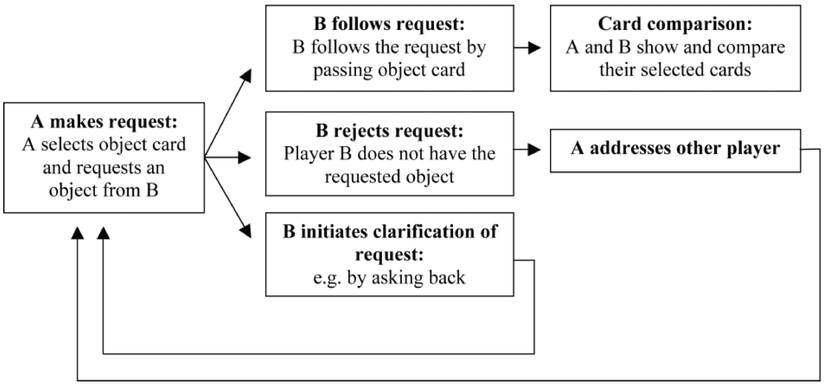


Figure 6.1: Object Request LAG. The diagram presents possible decisions that can be make during the game [Difrancesco et al., 2012].

The aim of the present study is therefore to extend the range of the rehabilitation focus that RGS originally provides to the rehabilitation of Broca's aphasics and to learn about the potential benefits of VR based language rehabilitation techniques.

6.2 Methods

In the present study we built a VR version of ILAT using the Rehabilitation Gaming System (RGS) in order to investigate the potential of VR based language rehabilitation methods. We conducted a double-case initial pilot study in order to test the system. Additionally, we evaluated the patients' language skills before and after the intervention

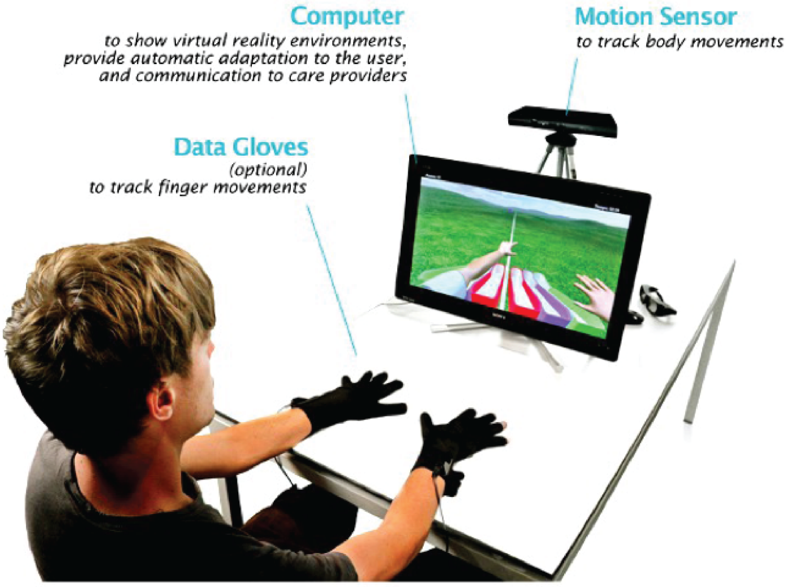


Figure 6.2: Rehabilitation Gaming System (RGS) setup. The subject works with his/her arms on a cut-out table facing a computer screen. The movements of the arms are tracked by the Kinect. The captured movements are mapped in real time to the movements of two virtual arms that mimic the movements of the user on the display.

6.2.1 System Description

The software of the present system is integrated within the environment of the RGS. The experimental setup consists of two standard personal computers (Vaio L Series All-in-One PC, Tokyo, Japan) with a 24" (61cm) Full HD touch screen, two motion sensing input devices (Xbox Kinect, Seattle, USA) and a networking system.

6.2.2 Virtual Scenario

The virtual environment was designed and developed using Unity3D game engine (Unity Technologies, San Francisco, USA). The motion tracking system (Xbox Kinect, Seattle, USA) captures and maps the movements of the users’ real arms onto those of the avatars. Consequently, during every session all users can continuously observe the movements of both their virtual arms’ and those of the opponent/s.

The training scenario takes place in a shared virtual room, to which the players connect via a local network. Each of the users can see three objects placed on the virtual table 6.3. In the beginning of the game one of the players is indicated to start the game. Consistent with the original request LAG scenario, his/her task is to first choose one object from those available on the table and then verbally request the same object from the opponent. In case of a successful communication the opponent should understand the request and pass the corresponding object. Such a sequence of events accounts for a successful communicative interaction, for which both players get a point.

The interaction with the virtual objects is based on delays. Thus in order to select or to pass an object, depending on the turn, players needs to place one hand over that object for 2-3 seconds. If the passed object matches the requested object they both appear in the basket which belongs to the player who started the round. At the same time the patients’ scores increase, two new objects appear on the table, and the turn changes. After such sequence of events, the second player is required to choose and request an object. The selection of objects is indicated by a short animation and a corresponding sound (e.g. a piano melody, in case of the piano; heartbeat in case of the heart; or footsteps in case of a shoe). We decided to incorporate sounds in order to provide additional associative cues assisting in recall and to help the patients in retrieving the words. All the objects light up whenever they are being pointed to. The purpose of the visual feedback is to enhance the salience of relevant objects and to ease the interaction with the system. Additionally, during the period of object selection an animated wall appears in the middle of the table. This prevents the opponent from seeing not only what object is being selected, but also possible indicative gestures. As soon as an object is selected, the wall disappears and the players can see the opponent’s virtual avatar. Instructive headings such as “It’s Your Turn”, “Well Done”, “Try Again”, or “Game Over” are displayed every time an event in the game changes (change of turn, collection of an object, failure, end of the game etc.). Since the system was tested in the Hospital Esperanza in Barcelona, Spain, the User Interface (UI) is written in Spanish.

Figure 6.3: The virtual scenario of Intensive Language Action Therapy in Virtual Reality from the first person perspective. The virtual objects are placed on the virtual table as well as in the baskets on the sides. The opponent is sitting at the other side of the table.

6.2.3 Setup

All the phases of the study took place in the Hospital de la Esperanza, in Barcelona, Spain. In a speech therapy ward, two computers were placed in front of each other so that the players could be close enough to efficiently communicate with one another. The seats were placed so that the patients could not see the opponent's real hands while selecting the virtual objects.

6.2.4 Protocol

Each of the patients participated in three phases of the study: a system evaluation phase, a training phase and an intervention phase. Both patients completed excerpts from Western Aphasia Battery (WAB) before and after the intervention, to be later analyzed. WAB is a standardized measure commonly used to assess the function of language which includes: Spontaneous speech, Auditory verbal comprehension and recognition, Sequential commands, Repetitions, Object naming, Word fluency Sentence completion Responsive speech and Reading comprehension of sentences and commands. After completing the evaluation, patients participated in twenty minutes training phase. During this phase a healthy player and a speech therapist were performing the virtual task. Meanwhile, the crucial parts of the game were being explained: system's startup, gaming rules and objectives. The next day the training phase took place (20 minutes). Two patients were asked to play against each other, and later they were interviewed about the usability of the system. Based on the foregoing evaluation slight changes were immediately incorporated with regards to the objects displacement within the VE. As the pre-phase period was completed the patients started the intervention which lasted for five days. Each of the patients played against the healthy player in the presence of a speech therapist. The role of both the healthy player and the therapist was to actively monitor the patients, keep track of their utterances, and adjust the velocity of the game, while the record of patients' successes and failures was stored by the system (RGS). Date, the session number, time, utterance type, quantity of failures, as well as scores from every session were continuously registered for further analysis. Moreover, all the sessions were recorded to extract Reaction Times (RTs) and to later analyze the data gathered from every session as suggested by the original study [Difrancesco et al., 2012]. RTs were measured from the moment when the patient selected the object to be requested from his/her opponent to when s/he fully uttered the correct name of the corresponding object. The game events, which included failures, names of the indicated objects, points, and the acts of selecting and passing the objects were continuously logged and stored after every session.

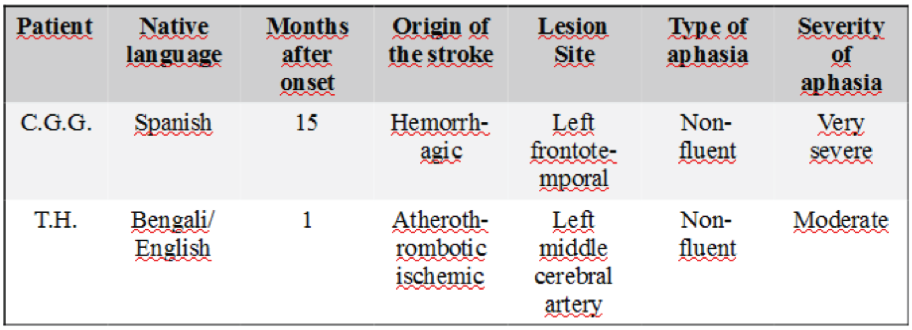
In order to further measure the potential change on a communication rating scale, the two patients as well as the speech therapist completed the Communicative Activity Log (CAL) before and after the intervention [Pulvermüller, 2005]. CAL is a qualitative tool to measure patients' amount and quality of communication in everyday life. CALs’ questions regarded the frequency with which patients would communicate in everyday life situations such as shopping, talking on the phone, answering/asking questions, and more. The questionnaire consisted of 18 questions to be answered on a 6-point Likert scale. The scale ranged from ‘never’ to ‘very frequently’. Additionally, after the period of the intervention, both patients were asked to complete a 16-item System Validation Questionnaire (SVQ) presented on a 7-point Likert scale.

6.2.5 Subjects

Two post- left hemispheric stroke patients 6.4, C.G.G. (man, aged 75, right- handed) and T.H. (woman, aged 52, left- handed) participated in the three experimental phases. Both subjects had normal vision and suffered from post stroke Broca’s aphasia. For the purpose of the present pilot study the exclusion criteria only partially followed the protocol introduced by Pulvermüller [Difrancesco et al., 2012]. We therefore first made sure that the two subjects did not suffer chronic heart disease or any related illnesses which makes the participation difficult. Secondly, the subjects could not suffer from any disease which would prevent from understanding the instructions of both the scenario of the LAGs and the interaction with the system itself. Therefore, the presence of impairments which affect perception as well as motor and neuropsychological functions such as deficits in motor planning (apraxia), vision, learning, memory or attention were accordingly evaluated and excluded.

6.3 Results

The present study was designed to investigate whether VR based language rehabilitation systems can trigger positive changes in the communicative behavior of post stroke Broca’s aphasia patients. We approach our aim by designing and testing ILAT in RGS system. Together the results presented here reinforce the notion that such novel techniques should be further



Patient	Native language	Months after onset	Origin of the stroke	Lesion Site	Type of aphasia	Severity of aphasia
C.G.G.	Spanish	15	Hemorrhagic	Left frontotemporal	Non-fluent	Very severe
T.H.	Bengali/ English	1	Atherothrombotic ischemic	Left middle cerebral artery	Non-fluent	Moderate

Figure 6.4: Clinical data about the patients who participated in the study (C.G.G. and T.H.)

investigated to better understand their efficiency and usability.

6.3.1 Clinical Evaluations: CAL and WAB

Results from CAL show improvements in all the evaluations in both patients (see Figure 6.5 *Left*). From the overall score of 90 points, the speech therapist assigned 43 points before and 47 points after the intervention to C.G.G. which means that the score increased by 9.3%. An increase was also reported in case of T.H. The score given by the therapist to T.H. prior to the intervention equaled to 19 points, and increased to 32 points after the treatment. The score increased by 68.4% after the intervention. Improved results from CAL were also observed in patients’ selfratings after the intervention. The score of C.G.G. increased by 14.7%, and the score of increased by T.H. by 27.9%.

The patients were asked to complete excerpts from the Western Aphasia Battery (WAB) before and after the intervention (see Protocol). The excerpts included the evaluation of “Spontaneous Speech” (20 points), “Auditory Verbal Comprehension” (200 points), “Repetition” (100 points), “Naming” (110 points) and “Reading” (100 points). Accordingly, the maximum score which could be achieved equaled 530 points. Both patients scored higher in the post evaluation than in the pre-test (see Figure 6.5 *Left*). Results from WAB prior to the intervention show an increase for C.G.G. and T.H. by 20.7% and 11.4% respectively. C.G.G. scored 232

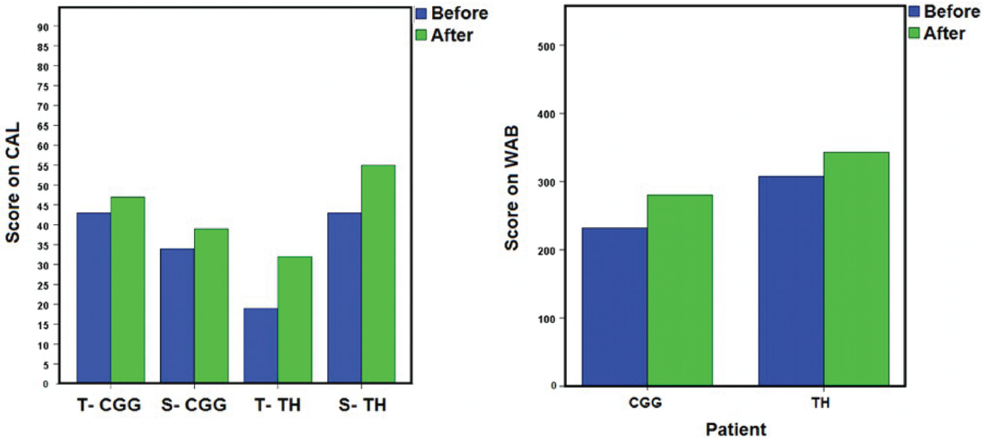


Figure 6.5: *Left:* Pre- and post- results from CAL for CGG and TH. Blue bar: results from the pre-evaluation, green bar: post-evaluation. T-CGG: therapist's evaluation of CGG T-T.H.: therapist's evaluation of T.H., S-CGG: self-evaluation of CGG, T-TH: selfevaluation of TH. *Right:* Results from pre- and post- WAB for C.G.G. and T.H. Blue bar: pre-evaluation score, green bar: post-evaluation score.

points and his result increased to 280 points after the intervention sessions. T.H. had 308 and 343 points before and after the intervention.

6.3.2 Data from the System

The mean RTs from every session was reported and compared throughout the five days of the intervention for C.G.G and T.H. (see Figure 6.6 *Left*). The reported data shows that the RT of C.G.G. decreased from 22 seconds (first day) to 10.75 seconds (fifth day), that is, by 51.1%. Similarly, in case of T.H. the RT also decreased from 22 (first day) seconds to 6.3 seconds (fifth day), that is, in 71.4%.

We define a failure as an event when a patient passes an object different than the one which was requested by the opponent. We believe that such behavior is caused by either confluent interaction with the system or a misunderstanding of the requested object. Over the period of the intervention the number of FTs in case of C.G.G. decreased by 62.5% and in case of T.H. it decreased by 55% (see Figure 6.6 *Middle*).

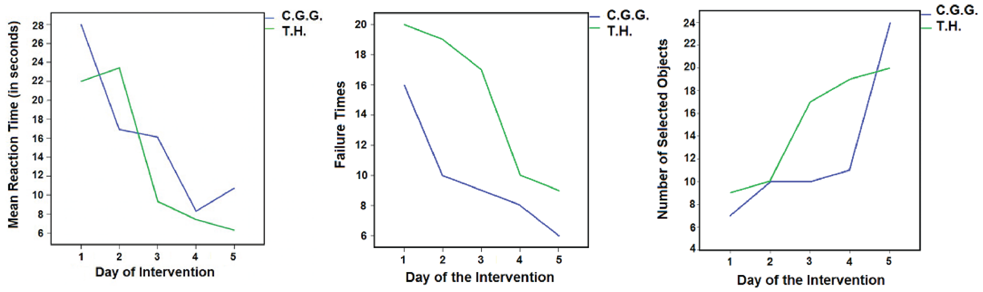


Figure 6.6: *Left:* The total scores obtained during every session by the two patients. Green: the score of T.H.; blue: the score of C.G.G. *Middle:* The number of Failure Times in every day of the intervention for the two patients.Failure times: the times when a patient gave an object different the one requested. Green: the number of FTs for T.H.; blue: the number of FTs for C.G.G.. *Right:* The number of objects selected during every session by the two patients. Green the number of objects selected by T.H; blue: the number of objects selected by C.G.G..

We considered the amount of objects selected representative for the fluent interaction with the system. In both patients the quantity of objects selected per session increased. C.G.G. selected 24 objects during the last session which was 3.8 times more than during the first day (see Figure 6.6 *Right*). The number of selected objects in case of T.H. was 2.2 times higher on the last day of the intervention. The patient selected 9 objects during the first and 20 objects during the last session.

The System Evaluation Questionnaire was distributed after the period of the intervention. It consisted of 16 questions which regarded patients' opinion on the system, its usability, functionality, design. The patients were to declare to which extend they agreed with a given statement on a 7-point Likert scale (see Protocol). The maximum score was 112 points from which C.G.G. scored 89 (79.5%) and T.H. 109 points (97.3%) respectively.

6.4 Conclusions and Discussion

In the present study we designed and developed a VR based language rehabilitation tool by integrating ILAT's request LAG to the RGS, and tested the system. The principal purpose of the initial pilot study was to validate the system and to learn whether a virtual adaptation of the Intensive Language Action Therapy into RGS can trigger positive changes in the linguistic behavior of Broca's aphasia patients. The gathered results suggest that both subjects learnt how to interact with the initial model of RGS-ILAT and they were satisfied with the system. Since Broca's aphasia patients do not suffer comprehension deficits [McNeil et al., 1990], FTs were mainly associated with the fluency and facility with which the users interacted with the system. After the intervention, the FTs in case of both C.G.G. and T.H. decreased. These results show that both patients got gradually acquainted with the system. The patients could interact more easily within the VR scenario, manipulate the objects and, as a consequence, play with more fluency. Additionally, the results from the SEQ were highly promising. C.G.G. and T.H. evaluated the system for 89 (79.5%) and 109 points (97.3%) respectively. The assessed attributes included facility, comprehensibility, fluency, effectiveness, the range of practiced vocabulary, entertainment and more. The questionnaire was also testing whether the patients felt entertained, motivated and satisfied while interacting with the system. To the statement "I would like to have the system at home" both patients strongly agreed. No further improvements were suggested by the participants. After having analyzed the overall results from the study, it can be argued that the approval and acceptance of the system was high. The reported positive changes in pre and post both WAB and LAG suggest that C.G.G. and T.H. improved their core language skills as well as communication skills. Since the RTs decreased in case of both patients over the period of the therapy, it may be concluded that both patients improved their language behavior within the 5-day intervention. Moreover, in both patients, the number of scored points in the game rose or remained unchanged through the sessions, which accounts for the increase in fluency.

One of the limitations of the present study was the lack of implementation of a precise hand-tracking system, which would allow for the simultaneous use of language and motor actions (e.g. holding the card while requesting the represented object), highly encouraged by ILAT. Moreover, we have not yet implemented the action-planning LAG, as suggested by the original protocol [Difrancesco et al., 2012]. To reliably compare the RGS-ILAT with the original therapy we need to include increasing number of objects that would amplify the range of the lexicon used, by introducing words of different frequency, minimal pairs, semantic categories, and multi-feature objects. Although present results might have been additionally influenced by other factors than the intervention, such as the natural recovery processes, motivation, or personal attitude, the positive outcome encourages us to further develop and test the system. Since only two subjects participated in the study and the period of the intervention was limited to five days, we are not yet able to compare our results to those of the original therapy To fully validate the RGS-ILAT system we will therefore conduct a follow-up study with an increased number of subjects and a higher intensity of the intervention, to be able to compare our results with those of a similar non-VR, and investigate whether the present system is more, less or equally effective. This will also shed the light on whether the positive rating of the proposed therapy is influenced by the novelty effect. Based on previously discussed results, the gathered feedback, as well as limitations we will provide the present ILAT-RGS with substantive technological advancements and evaluate the system in order to better understand the potential of the virtual reality based language rehabilitation therapies. The reason for proceeding with the enhancements of the RGS-ILAT system is to provide aphasia patients with an effective language rehabilitation tool which could be utilized as an additional reinforcement to the conventional therapy, or its continuation.

Chapter 7

AUGMENTED DYADIC THERAPY BOOSTS RECOVERY OF LANGUAGE FUNCTION IN PATIENTS WITH NONFLUENT APHASIA: A RANDOMISED CONTROLLED TRIAL

This chapter is based on:
Grechuta, K., Ballester, B. R., Munné, R. E., Bernal, T. U., Hervás, B. M., Mohr, B., Pulvermuller, F., San Segundo, R., & Verschure, P. (2019) Augmented embodied dyadic therapy boosts recovery in patients with nonfluent aphasia: a randomised controlled trial. *Stroke.*

Evidence suggests that therapy can be effective in recovering from aphasia, provided that it consists of socially-embedded, intensive training of behaviorally relevant tasks. However, the resources of health-care sys-

tems are often too limited to provide such treatment at sufficient dosage. Hence, there is a need for evidence-based, cost-effective rehabilitation methods. Here, we asked whether virtual reality-based treatment grounded in the principles of use-dependent learning, behavioral relevance, and intensity positively impacts recovery from nonfluent aphasia. Seventeen patients with chronic nonfluent aphasia underwent intensive therapy in a randomized, controlled, parallel-group trial. Participants were assigned to the Control Group (CG, N=8) receiving standard treatment or to the Experimental Group (EG, N=9) receiving augmented embodied therapy with the Rehabilitation Gaming System for aphasia (RGSa). All RGSa sessions were supervised by an assistant who monitored the patients but did not offer any elements of standard therapy. Both interventions were matched for intensity and materials. Our results revealed that at the end of the treatment both groups significantly improved on the primary outcome measure (Boston Diagnostic Aphasia Examination, CG: p=.04; EG: p=.01), and the secondary outcome measure (lexical access: VocabT, CG: p=.01; EG: p=.007). However, only the RGSa group improved on the Communicative Aphasia Log (p=.01). The follow-up assessment (week 16) demonstrated that while both groups retained vocabulary-related changes (CG: p=.01; EG: p=.007), only the RGSa group showed therapy-induced improvements in language (p=.01) and communication (p=.05). Our results demonstrate the effectiveness of RGSa for improving language and communication in patients with chronic aphasia suggesting that current challenges faced by the health-care system in the treatment of stroke might be effectively addressed by augmenting traditional therapy with computer-based methods.

7.1 Introduction

20% of post-stroke patients display nonfluent aphasia at the chronic stage ($\geq$ 6 months post stroke). Affected individuals may experience changes in language processing and learned nonuse [Pulvermüller and Berthier, 2008], resulting in social exclusion, depression, a compromised quality of life, as well as limited language recovery [Hilari et al., 2012]. Although clinical evidence suggests that therapy can facilitate rehabilitation, provided that it consists of socially-embedded, intensive training of behaviorally relevant goal-oriented tasks [Pulvermüller and Berthier, 2008, Berthier and Pulvermüller, 2011] the resources of most health-care systems are too limited to promote such methods at sufficient dosage [Katz et al., 2000, Gunning et al., 2017]. Hence, there is a need for evidence-based, cost-effective rehabilitation techniques for improving the condition of people with aphasia and maximizing their self-efficacy.

Several studies suggest that computer-based methods could be beneficial for providing rehabilitation for post-stroke aphasia patients [Lavoie et al., 2017, Zheng et al., 2016]. Importantly, however, evidence from Randomized Controlled Trials (RCT) supporting the implementation of such treatments in the clinic is still required. Here, we pose the question of whether embedding aphasia rehabilitation in the context of embodied peer to peer interaction grounded in the principles of use-dependent learning, behavioral relevance, and intensity [Kleim and Jones, 2008, Stahl et al., 2016]positively impacts recovery. To the best of our knowledge, we performed the first RCT investigating the effectiveness of this approach for persisting nonfluent aphasia by using a Virtual Reality (VR)-based rehabilitative technique. The proposed therapy called the Rehabilitation Gaming System for aphasia (RGSa) provides lexical and syntactic training in a multimodal, goal-oriented manner within a context of dyadic peer-interaction [Kleim and Jones, 2008, Stahl et al., 2016]. We hypothesized that RGSa training would lead to a comparable recovery of language functions as a standard Speech and Language Therapy (SLT).

7.2 Methods

The data that support the findings of this study are available from the corresponding author upon request. The RCT was approved by the local ethics committee and registered on ClinicalTrials.gov (identifier: NCT02928822).

Seventeen patients 7.1 with chronic nonfluent aphasia provided their consent and participated in a parallel-group RCT. We used computer-generated stratified randomization to assign the participants to the Experimental (EG, N=9) or Control Group (CG, N=8). Clinical and demographic sample characteristics are presented in the Supplement (Table1).

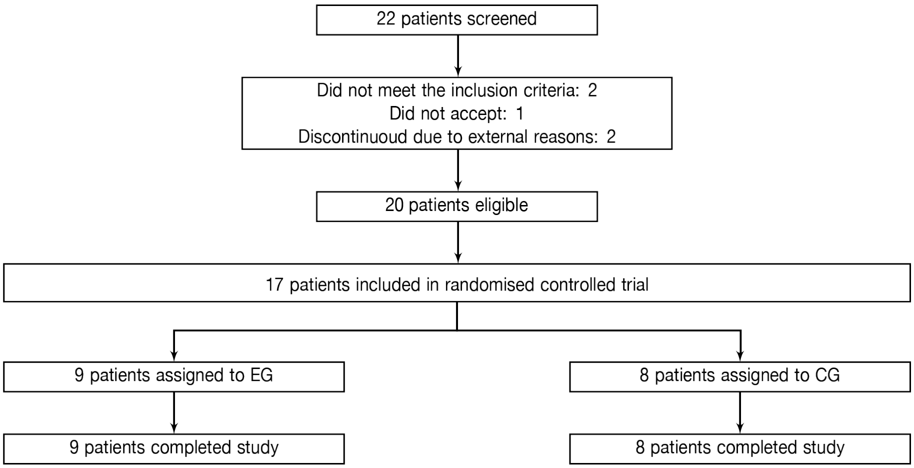


Figure 7.1: Consolidated standards of Reporting Trials flow diagram.

The RGSa (Figure 7.2 A) had a form of dyadic peer to peer language-training protocol inspired by Intensive Language-Action Therapy (ILAT) [Difrancesco et al., 2012]. Two patients sat in front of each other facing their respective screens. They interacted by performing planar arm movements which were tracked and mapped onto the avatars’ upper-limbs providing embodiment and allowing the interaction with virtual objects (Figure 7.2 B-C). The paradigm required engagement in everyday-like communication acts by requesting objects, or handing them over when requested by the other player [Difrancesco et al., 2012]. There was a set of three objects simultaneously available for selection. Object selection for both request and response required the players to place the hand over the target object for 3 seconds.

The interaction was based on turns (Figure 7.2 D), and the goal for each patient and each session was to collect 36 objects. The materials consisted of 120 three-dimensional objects (Figure 7.2 B-C). To promote the activation of the language network the RGSa protocol allowed using gestures and self-cueing strategies which accompanied but did not substitute verbal communication [Pulvermüller and Fadiga, 2010]. The RGSa sessions were supervised by a therapy assistant who monitored possible technological or communication difficulties. Importantly, the assistant did not offer any additional services.

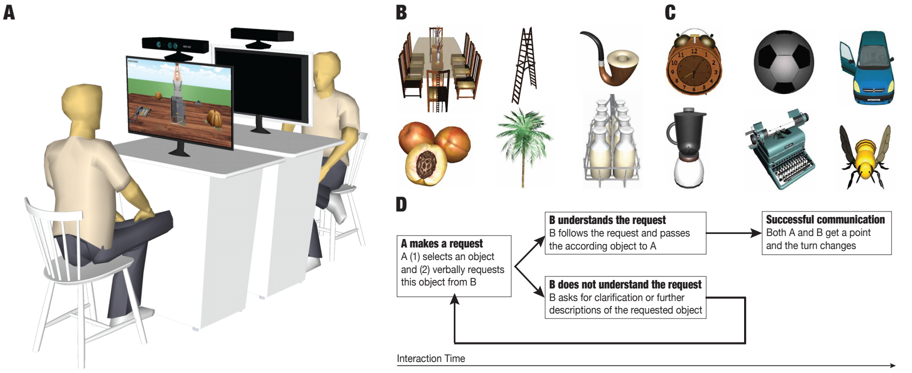


Figure 7.2: Setup and materials. Therapeutic setting (A). Example of materials (B-C). Items without (B) and with (C) semantically related sounds. Illustration of the dynamics of the RGSa interaction, possible moves and speech-acts (D).

In the Control Group (CG), patients received standard SLT targeting specific linguistic deficits in a therapist-patient setting. The therapy aimed at training naming, repetition, spelling, and articulation. The intensity and frequency of the intervention in both groups were matched: eight weeks, five days per week. The duration of each session was 30-40min. Similarly, both groups used the same training materials: EG-122 3D objects (Figure 7.2 B-C); CG-122 images presented in the form of cards.

The primary outcome measure was the performance on chosen subtests from the Boston Diagnostic Aphasia Examination (BDAE) which measures language function in people with language disorders. The secondary outcome measures included the Communicative Activity Log (CAL) evaluating communicative frequency and effectiveness in everyday life, the Vocabulary Test (VocabT) assessing the lexical access and verbal execution of the trained stimuli, Fugl-Meyer Upper Extremity Scale (FMA-UE) measuring the motor impairment of the hemiplegic arm, and the Interaction

Times (ITs, Figure2D) extracted from the RGSa system. The corresponding data were collected at baseline (T0), week 4 (T1), week 8 (end of intervention; T2), and the follow-up (week 16; T3). During the follow-up period, patients could not continue SLT or receive any therapy from third parties. Detailed information about the methods as well as the outcome measures is available in the Supplemental material.

We used the Friedman test to compare the overall effect of the treatments within groups and the Wilcoxon signed-rank test for post-hoc statistical analyses. The Mann-Whitney U-test was performed to identify between-group differences.

7.3 Results

The homogeneity of the groups at baseline was confirmed for all measures (Supplemental Table2).

Based on BDAE (Figure 7.3 A, Supplemental Table3 and Table4), an increase in language performance was found for both groups ($N = 17, p < .001$). Specifically, we observed changes in T2-T0 and T3-T0 ($p = .001 and p = .002$ respectively). The within-group analysis showed differences for both CG ($p = .04$) and EG ($p = .006$). The post-hoc analysis for EG demonstrated changes at T2-T0 ($p = .01$), T3-T0 ($p = .01$), and T3-T1 ($p = .01$). For CG, we found changes at T2-T0 ($p = .04$), and T2-T1 ($p = .04$). No differences in BDAE changes were found between groups. CAL analysis (Figure 7.3 B, Supplemental Table3) showed improvement for both groups ($p = .01$) at each time-step ($T1 : p = .01; T2 : p = .002; T3 : p = .009$). The within-group analyses yielded changes in EG ($p = .02$), in particular, an increase at T1-T0, T2-T0, and T3-T0 ($p = .01, p = .01, p = .05$) as well as T2-T1 ($p = .01$). However, we found neither effects for CG nor significant differences between the groups.

The VocabT analysis (Figure 7.3 C, Supplemental Table3) revealed an improvement at the whole-group level ($N = 17; p < .001$) at each time step ($T1 : p < .001; T2 : p < .001; T3 : p < .001$). The within-group analysis demonstrated differences for EG ($p < .001$) and CG ($p = .003$). For EG, the scores differed between T0 and T1, T2, and T3 ($p = .01, p = .007, p = .007$), between T1 and T2 ($p = .007$), and between T1 and T3 ($p = .007$). For CG, we found changes from T0 at T1, T2, and T3 ($p = .01, p = .01, p = .01$), from T1 at T2 ($p = .01$) and T3 ($p = .01$), and from T2 at T3 ($p = .04$). Additionally, the between-group analysis showed a significant difference between EG and CG at T1 ($p = .02$).

Based on the FMA-UE (Supplemental Table3), we found changes for the two groups at T1 ($p = .04$) and T2 ($p = .02$). The analysis yielded an effect of time in EG at T1 ($p = .04$) and T2 ($p = .04$). No differences were found for CG (p=.6). We further report differences at T1-T0, T2-T0, and T3-T0 between the groups ($p < .009, p < .04, p < .03$).

Finally, ITs decreased for all patients in EG over the therapy interval. In particular, linear regression revealed a relationship between the averaged ITs and the therapy days ($R = -0.38, p = .01$). Moreover, the Pearson correlation yielded a significant relationship between ITs and VocabT ($R = -0.46, p < .001$).

7.4 Discussion

We explored the effects of RGSa on functional recovery of language and communication in patients with nonfluent aphasia relative to standard SLT. Our results revealed that, immediately after treatment, both groups improved in terms of speech production, auditory comprehension, communicative effectiveness in everyday life, and lexical access. At the follow-up, both groups retained the vocabulary-related changes. However, only the RGSa maintained the language improvements.

First, the experimental design controlled for the influence of training intensity and duration, clinical setting, therapy materials, as well as the number of utterances. However, the two groups differed regarding the training components (embodied, sensorimotor, goal-oriented training vs. linguistic, impairment-focused), content (verbal communication vs. naming and repetition), and the nature of the interaction (peer-peer

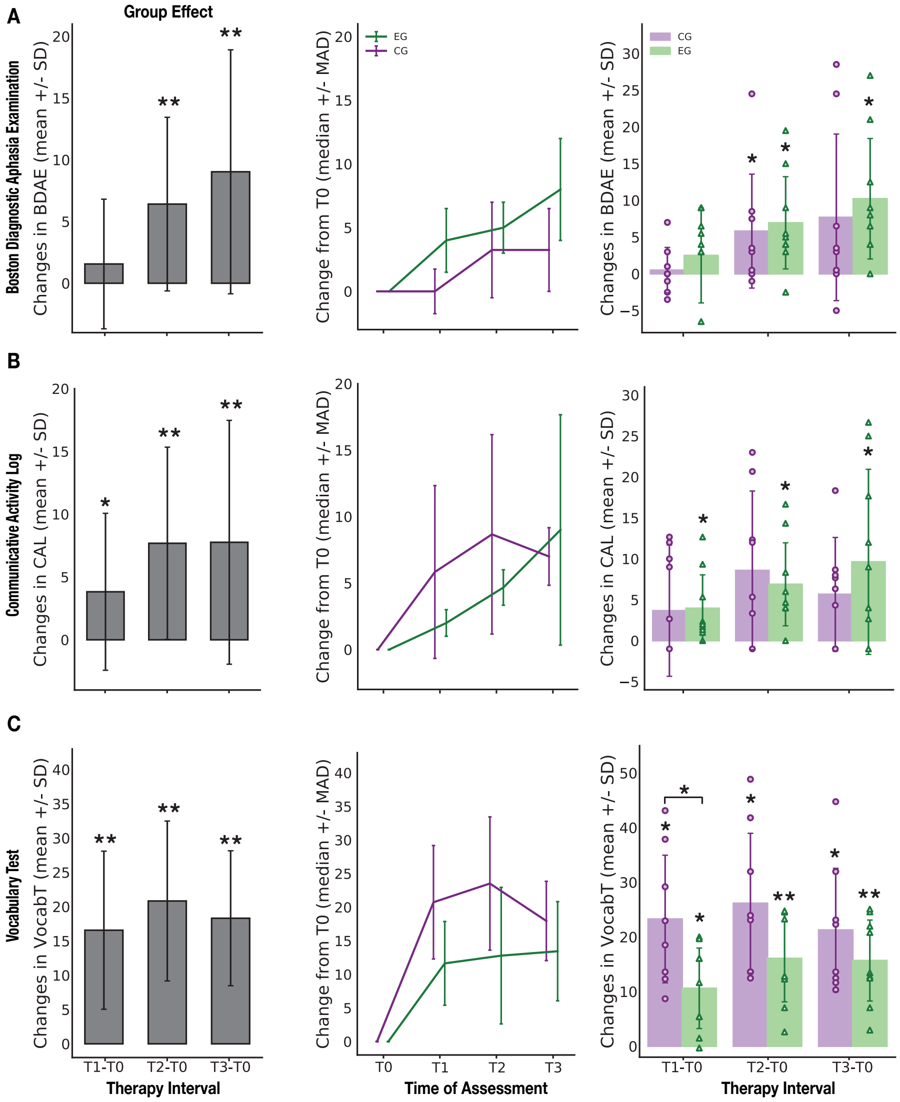


Figure 7.3: Clinical outcomes. BDAE (A), CAL (B), and VocabT (C). Left columns: whole-group effect (N=17). Central and right columns: changes from baseline at each time step for EG and CG. Triangles and circles: individual-patient data. Asterisks (*P¡0.05, **P¡0.01): Wilcoxon signed-rank and Mann-Whitney U-test.

vs. therapist-patient). Additionally, in the RGSa group, the communication consisted of the contextualized peer to peer training and, at times, exchanges with the therapy assistant. Despite the methodological heterogeneity, however, both methods were beneficial for language improvement and retention of changes in the lexical access. Notably, CG performed significantly better than RGSa on VocabT at T1, and RGSa outperformed CG on CAL at each assessment point. These results may suggest that while the SLT promoted lexical access and verbal execution of the target stimuli, RGSa possibly induced higher frequencies of language use. Moreover, only the RGSa group showed significant maintenance of those improvements. Future studies should systematically investigate the contribution of the specific training components of the RGSa on functional and structural recovery as well as their short and long-term effects with a higher sample-size.

Second, RGSa allowed the use of both arms and the execution of gestures serving self-cueing strategies while preventing compensatory mechanisms. This design principle is grounded on the evidence of bidirectional connections between the motor and language systems [Pulvermüller and Fadiga, 2010]. Specifically, it supports the notion that enhanced neuronal activity during language processing through simultaneous execution of goal-directed and language-associated gestures might facilitate language performance, and vice-versa. Indeed, we found that EG, but not CG, showed significant differences in FMA-UE during the intervention. Although these differences do not seem to reach clinical relevance [Page et al., 2012], they show a positive trend suggesting that RGSa might also promote the recovery of motor function. Based on the same principle, future studies shall investigate the reverse effect, specifically, the integration of action-embedded language use in the recovery of motor disorders.

Overall, present results emphasize the potential of RGSa for the improvement and long-term stability of the language training effects. From a health-care perspective, RGSa might be integrated into the clinical practice allowing people with aphasia access to continuous and self-paced training at the chronic stage of the disease.

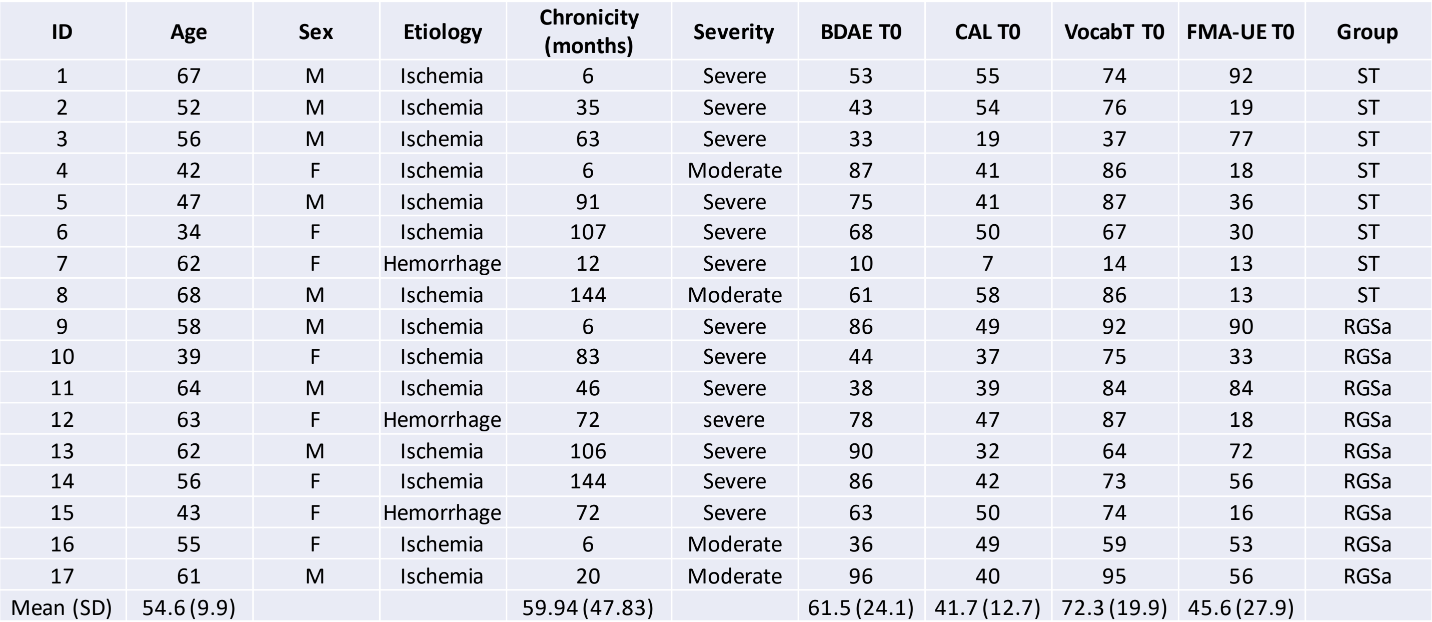
7.5 Supplement

7.5.1 Supplemental Methods

We conducted a parallel-group, longitudinal Randomized Controlled Trial (RCT), which took place at the l'Hospital Universitari Joan XXIII de Tarragona (Spain) from 2016 to 2018. This study builds on preliminary work exploring the possible effects, feasibility, and safety of the RGSa system [Grechuta et al., 2016a].

Participants

Inclusion criteria were the following: (1) moderate or severe stages of non-fluent aphasia as identified by a standard screening tool [Peña-Casanova, 1990], (2) more than 5 months post first-ever stroke, (3) right-handedness assessed by the Edinburgh Handedness Inventory [Oldfield, 1971], (4) age between 25 and 85 years old. We excluded patients with severe and untreated forms of cognitive disorders (assessed by the Mini-Mental State Examination) and motor impairments (assessed using Fugl-Meyer Assessment Upper Extremity) which could adversely affect interaction with the RGSa. Additionally, we excluded individuals who two years prior to the enrollment participated in alternative intensive interventions. All patients suffered a single left-hemispheric stroke affecting frontotemporal and parietal cortical areas, as evidenced by MRI or CT-scans. Patients presented moderate-to-severe aphasia. Clinical and demographic variables for all the participants are provided in supplemental Table1. The principal investigator and a speech and language therapist administered the associated recruitment procedures. Potential participants were contacted via phone and invited for a screening session. From the 22 patients screened, 2 did not meet the inclusion criteria, 1 did not accept to participate in the study, and 2 withdrew due to difficulties in transportation from home to the center (N=1, CG) and health issues unrelated to the study (N=1, EG). Consequently, 20 individuals were recruited and screened for eligibility (Figure1).



ID	Age	Sex	Etiology	Chronicity (months)	Severity	BDAE T0	CAL T0	VocabT T0	FMA-UE T0	Group
1	67	M	Ischemia	6	Severe	53	55	74	92	ST
2	52	M	Ischemia	35	Severe	43	54	76	19	ST
3	56	M	Ischemia	63	Severe	33	19	37	77	ST
4	42	F	Ischemia	6	Moderate	87	41	86	18	ST
5	47	M	Ischemia	91	Severe	75	41	87	36	ST
6	34	F	Ischemia	107	Severe	68	50	67	30	ST
7	62	F	Hemorrhage	12	Severe	10	7	14	13	ST
8	68	M	Ischemia	144	Moderate	61	58	86	13	ST
9	58	M	Ischemia	6	Severe	86	49	92	90	RGSa
10	39	F	Ischemia	83	Severe	44	37	75	33	RGSa
11	64	M	Ischemia	46	Severe	38	39	84	84	RGSa
12	63	F	Hemorrhage	72	severe	78	47	87	18	RGSa
13	62	M	Ischemia	106	Severe	90	32	64	72	RGSa
14	56	F	Ischemia	144	Severe	86	42	73	56	RGSa
15	43	F	Hemorrhage	72	Severe	63	50	74	16	RGSa
16	55	F	Ischemia	6	Moderate	36	49	59	53	RGSa
17	61	M	Ischemia	20	Moderate	96	40	95	56	RGSa
Mean (SD)	54.6 (9.9)			59.94 (47.83)		61.5 (24.1)	41.7 (12.7)	72.3 (19.9)	45.6 (27.9)	

Figure 7.4: Clinical and Sociodemographic Patient Characteristics

Randomization and masking

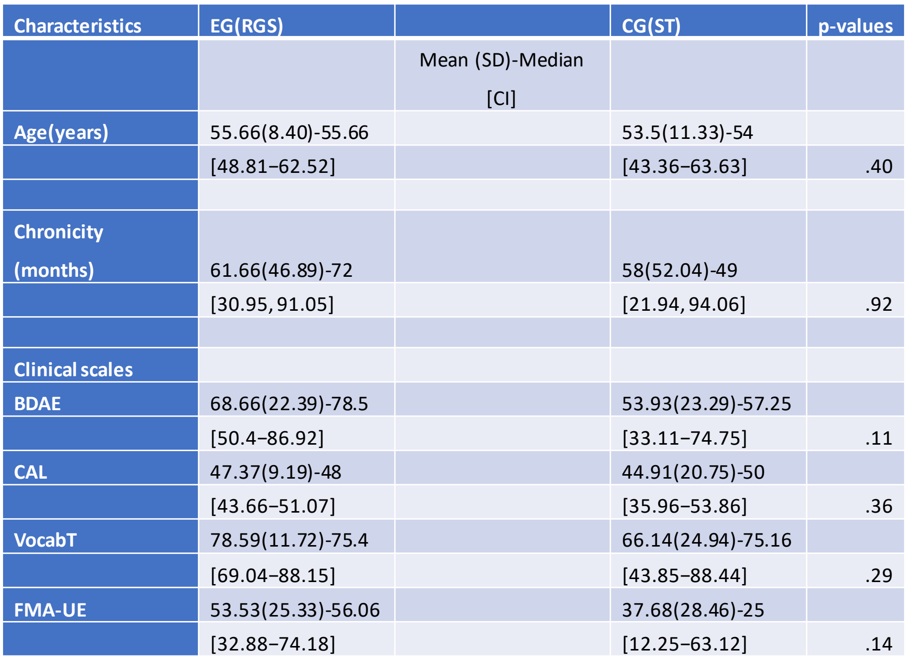
We used computer-generated stratified randomization to assign the participants to either the Experimental (EG) or the Control Group (CG). Consequently, 9 participants were assigned to EG and 8 to CG. Demographical data, as well as clinical records at baseline evaluation, are shown in the supplemental Table1 and the group characteristics in the supplemental Table2.

Clinical evaluations were conducted by four assessors including a speech and language therapist and three physiatrists. Each assessor underwent a competency training prior to the assessments. To avoid biases in the assessment, all data were cross-checked by an evaluator who was not involved in data collection. The assessors were instructed not to disclose to the patients neither the purpose of the study nor the perceived linguistic changes.

Outcome measures

Corresponding data were collected at the treatment site at baseline (T0), week 4 (T1), week 8 (the end of the intervention; T2), and the follow-up (week 16; T3). After the end of the trial, patients were not allowed to continue the SLT during the follow-up period and did not receive any therapy from third parties. Baseline characteristics included demographic data, stroke classification as well as the clinical metrics (Supplement, Table1).

The primary outcome measure was the performance on the chosen subtests from the Boston Diagnostic Aphasia Examination (BDAE) [García-Albea et al., 1996]. In particular, we assessed subtests related to auditory comprehension, automatic speech, repetition, and naming. Since both types of treatment focused on expression and comprehension of spoken, rather than written, language, we did not evaluate items related to reading and writing. The BDAE subtests scores were summed up and normalized to obtain the total score for each patient. We considered the primary endpoint as the change between T0 and T3 for each group. The secondary endpoints included T2-T0 and T1-T0.



Characteristics	EG(RGS)		CG(ST)	p-values
		Mean (SD)-Median [CI]		
Age(years)	55.66(8.40)-55.66		53.5(11.33)-54	
	[48.81–62.52]		[43.36–63.63]	.40
Chronicity (months)	61.66(46.89)-72		58(52.04)-49	
	[30.95, 91.05]		[21.94, 94.06]	.92
Clinical scales				
BDAE	68.66(22.39)-78.5		53.93(23.29)-57.25	
	[50.4–86.92]		[33.11–74.75]	.11
CAL	47.37(9.19)-48		44.91(20.75)-50	
	[43.66–51.07]		[35.96–53.86]	.36
VocabT	78.59(11.72)-75.4		66.14(24.94)-75.16	
	[69.04–88.15]		[43.85–88.44]	.29
FMA-UE	53.53(25.33)-56.06		37.68(28.46)-25	
	[32.88–74.18]		[12.25–63.12]	.14

Figure 7.5: Patient’s characteristics at baseline (n=17)

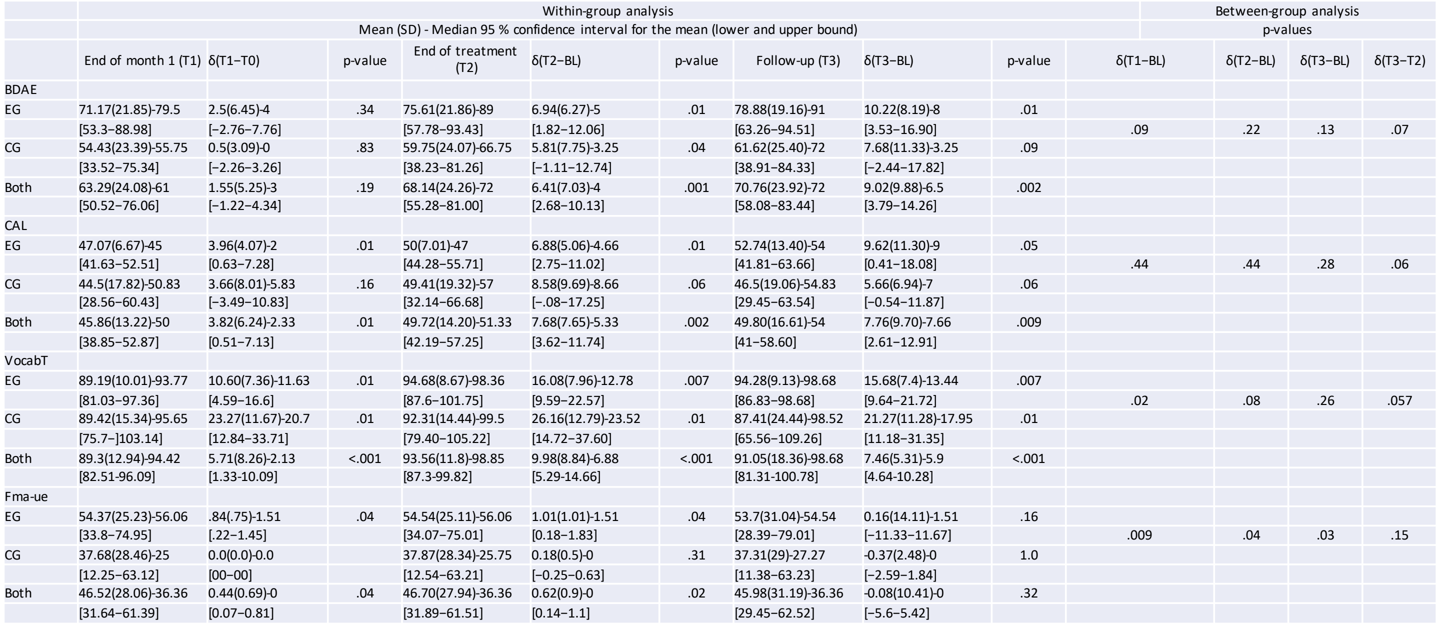
To evaluate changes in communicative skills, a therapist blind with regards to group assignment, caregivers, and patients completed the Communicative Activity Log (CAL) [Pulvermüller et al., 2001b]. CAL is a quantitative tool measuring the amount and quality of everyday communication during common activities of daily life (see below the full list of items). The questionnaire consists of 18 items (the full list of questionnaire items is listed below) assessing the frequency of communication using a Likert scale ranging from 0: “never” to 5: “very frequently”. For the analysis, to counteract potential biases in the assessment, we averaged the scores from the three evaluators for each participant, and we considered T3-T0, T2-T0, and T1-T0 as secondary outcomes.

Another secondary outcome included learning of the target stimuli as measured by a Vocabulary Test (VocabT) which included all the trained items (N=120). For each word, patients could score a maximum of 5 points

(0: no verbal utterance, 1: utterance followed by full phonetic priming, 2: utterance followed by priming of the initial phoneme, 3: utterance followed by full silent orofacial hint, 4: utterance followed by silent orofacial hint of the first phoneme, 5: utterance followed by no hint). The test was administered five times over the intervention period (at randomization, week 2, 4, 6, 8), and at the follow-up. For the analysis of the improvement in the lexical access, we considered the primary endpoint as T3-T0, and the secondary ones included T2-T0 and T1-T0. The intermediate timesteps (evaluation at week 4 and 6) served as a reference for the correlation with the Interaction Times (ITs) extracted from the system (Figure2D).

We conducted the Fugl-Meyer Upper Extremity Scale (FMA-UE) of the hemiplegic right arm to evaluate potential changes related to the motor function. The assessment was conducted at T0, T1, T2, and T3. The primary endpoint was T2-T0 while the secondary endpoints were T3-T0 and T1-T0.

Finally, we extracted and computed Interaction Times (ITs, Figure 2D) for all stimuli and therapy sessions. We defined an IT as the time interval between the selection and collection of an object. Each IT included word retrieval, articulation of the request, comprehension of the target word and the motor response of the opponent.



	Within-group analysis									Between-group analysis			
	Mean (SD) - Median 95 % confidence interval for the mean (lower and upper bound)									p-values			
	End of month 1 (T1)	δ(T1−T0)	p-value	End of treatment (T2)	δ(T2−BL)	p-value	Follow-up (T3)	δ(T3−BL)	p-value	δ(T1−BL)	δ(T2−BL)	δ(T3−BL)	δ(T3−T2)
BDAE													
EG	71.17(21.85)-79.5	2.5(6.45)-4	.34	75.61(21.86)-89	6.94(6.27)-5	.01	78.88(19.16)-91	10.22(8.19)-8	.01				
	[53.3−88.98]	[−2.76−7.76]		[57.78−93.43]	[1.82−12.06]		[63.26−94.51]	[3.53−16.90]		.09	.22	.13	.07
CG	54.43(23.39)-55.75	0.5(3.09)-0	.83	59.75(24.07)-66.75	5.81(7.75)-3.25	.04	61.62(25.40)-72	7.68(11.33)-3.25	.09				
	[33.52−75.34]	[−2.26−3.26]		[38.23−81.26]	[−1.11−12.74]		[38.91−84.33]	[−2.44−17.82]					
Both	63.29(24.08)-61	1.55(5.25)-3	.19	68.14(24.26)-72	6.41(7.03)-4	.001	70.76(23.92)-72	9.02(9.88)-6.5	.002				
	[50.52−76.06]	[−1.22−4.34]		[55.28−81.00]	[2.68−10.13]		[58.08−83.44]	[3.79−14.26]					
CAL													
EG	47.07(6.67)-45	3.96(4.07)-2	.01	50(7.01)-47	6.88(5.06)-4.66	.01	52.74(13.40)-54	9.62(11.30)-9	.05				
	[41.63−52.51]	[0.63−7.28]		[44.28−55.71]	[2.75−11.02]		[41.81−63.66]	[0.41−18.08]		.44	.44	.28	.06
CG	44.5(17.82)-50.83	3.66(8.01)-5.83	.16	49.41(19.32)-57	8.58(9.69)-8.66	.06	46.5(19.06)-54.83	5.66(6.94)-7	.06				
	[28.56−60.43]	[−3.49−10.83]		[32.14−66.68]	[−.08−17.25]		[29.45−63.54]	[−0.54−11.87]					
Both	45.86(13.22)-50	3.82(6.24)-2.33	.01	49.72(14.20)-51.33	7.68(7.65)-5.33	.002	49.80(16.61)-54	7.76(9.70)-7.66	.009				
	[38.85−52.87]	[0.51−7.13]		[42.19−57.25]	[3.62−11.74]		[41−58.60]	[2.61−12.91]					
VocabT													
EG	89.19(10.01)-93.77	10.60(7.36)-11.63	.01	94.68(8.67)-98.36	16.08(7.96)-12.78	.007	94.28(9.13)-98.68	15.68(7.4)-13.44	.007				
	[81.03−97.36]	[4.59−16.6]		[87.6−101.75]	[9.59−22.57]		[86.83−98.68]	[9.64−21.72]		.02	.08	.26	.057
CG	89.42(15.34)-95.65	23.27(11.67)-20.7	.01	92.31(14.44)-99.5	26.16(12.79)-23.52	.01	87.41(24.44)-98.52	21.27(11.28)-17.95	.01				
	[75.7−]103.14]	[12.84−33.71]		[79.40−105.22]	[14.72−37.60]		[65.56−109.26]	[11.18−31.35]					
Both	89.3(12.94)-94.42	5.71(8.26)-2.13	<.001	93.56(11.8)-98.85	9.98(8.84)-6.88	<.001	91.05(18.36)-98.68	7.46(5.31)-5.9	<.001				
	[82.51-96.09]	[1.33-10.09]		[87.3-99.82]	[5.29-14.66]		[81.31-100.78]	[4.64-10.28]					
Fma-ue													
EG	54.37(25.23)-56.06	.84(.75)-1.51	.04	54.54(25.11)-56.06	1.01(1.01)-1.51	.04	53.7(31.04)-54.54	0.16(14.11)-1.51	.16				
	[33.8−74.95]	[.22−1.45]		[34.07−75.01]	[0.18−1.83]		[28.39−79.01]	[−11.33−11.67]		.009	.04	.03	.15
CG	37.68(28.46)-25	0.0(0.0)-0.0		37.87(28.34)-25.75	0.18(0.5)-0	.31	37.31(29)-27.27	-0.37(2.48)-0	1.0				
	[12.25−63.12]	[00−00]		[12.54−63.21]	[−0.25−0.63]		[11.38−63.23]	[−2.59−1.84]					
Both	46.52(28.06)-36.36	0.44(0.69)-0	.04	46.70(27.94)-36.36	0.62(0.9)-0	.02	45.98(31.19)-36.36	-0.08(10.41)-0	.32				
	[31.64−61.39]	[0.07−0.81]		[31.89−61.51]	[0.14−1.1]		[29.45−62.52]	[−5.6−5.42]					

Figure 7.6: Outcome measures at month 1, at the end of treatment, and follow-up. Bold values indicate significant differences (p¡.05). P-values for within-group analysis were obtained with Wilcoxon signed-rank test while p-values for between-group analysis were obtained with Wilcoxon rank-sum respectively.

Details of the intervention: setup, protocol, materials, and procedures

The proposed augmented dyadic therapy was integrated within the VR-based rehabilitation environment for motor and cognitive deficits, the so-called Rehabilitation Gaming System (RGS) [Grechuta et al., 2016a, Cameirão et al., 2010]. RGS for aphasia (RGSa) had a form of an interactive language-training protocol inspired by the Intensive Language-Action Therapy (ILAT) [Difrancesco et al., 2012].

The experimental setup (Figure 2A) consisted of two desktop computers (Vaio, Japan) connected via a local area network, two motion sensing input devices (Kinect2, Microsoft, USA) and two headsets (EX-01 BluetoothR, Gioteck, Canada). For the purpose of this RCT, during each session, patients sat in front of each other at a cut-out table facing their respective computer screens. Patients interacted by performing planar arm movements over a tabletop. All movements were continuously tracked by the sensor and mapped onto the avatars’ upper-limbs allowing the interaction with the virtual objects (Figure 2B-C). Each patient was provided with a single-ear wireless headset through which they heard instructions, feedback, and cues from the system. Patients could hear each other well and communicate without the use of technology. To prevent from novelty effects, prior to the intervention, all participants were familiarized with the interface and the dynamics of the game. All the sessions were supervised by a therapy assistant who received training on the use of the RGSa and whose role was to monitor the patients throughout the intervention and supervise them when a trial could not be completed independently. Importantly, the assistant did not offer any elements of standard therapy.

The RGSa paradigm required two patients to engage in everyday-like communication acts by requesting objects, or giving them when requested [Grechuta et al., 2016a, Difrancesco et al., 2012]. In particular, the protocol partially reflected the dynamics of ILAT’s object request Language Action Game (LAG) (Figure 2D). There was a set of three identical objects simultaneously available for selection, and both sets were always visible for both players. At the beginning of each session, one of the patients was randomly assigned to commence the game. Each trial consisted of three steps (Figure 2D): (1) player chooses the desired object by reaching towards it (this movement may involve both elbow extension and shoulder flexion), (2) player verbally request it, and (3) the opponent understands the request and reaches for the corresponding object. Such a sequence of events accounted for a successful communication after which both patients received a point, heard the correct pronunciation of the target word through headphones and the turn changed. In case the opponent did not understand the request, it needed to be repeated or clarified until the opponent understands it (Figure 2D). The goal for each patient and each session was to collect 36 objects.

The materials consisted of 120 three-dimensional objects (Figure 2B-C). To ensure that the stimuli were visually unambiguous and likely to elicit target stimuli, they were first evaluated by healthy participants and clinicians involved in the trial. For each pair of patients, the stimuli had the same difficulty and were presented in a pseudo-randomized order, counterbalanced within each week. All items were categorized according to (1) the frequency of occurrence in Spanish language including: high (“clock”, “table”), middle (“milk”, “ball”), and low frequency words (“pipe”, “palm tree”); (2) semantic category (i.e. food, instruments, animals); (3) complexity (simple: “car”, Spanish “coche” and complex: “ typewriter”, Spanish “ máquina de escribir”); and (4) phonemic similarity (i.e. minimal pairs: “bed” and “house”, Spanish “casa” and “cama”). The target words were matched for the syllable length and semantic field. Furthermore, they were classified into two categories: those which do not have semantically related sounds and in 50% of the trials underwent “Silent Visuomotor Cueing” [Grechuta et al., 2017a] (SVC, N=60, e.g. “bench”, “peach”, “milk”), and those which do have semantically related sounds and in 50% of the trials underwent “Semantic Auditory Cueing” (SAC, N=60, e.g. “telephone”, “car”, “typewriter “). SAC consisted of playing a representative sound of the object selected for a request and SVC consisted in displaying a silent video representing the correct pronunciation of the target word. Both cues were provided immediately after the object selection. The full list of the therapy items is provided below.

Differently from ILAT, in RGSa, both players (A and B, Figure 2D) were provided the same set of objects. Therefore, player B could either follow the request or ask for clarifications, but they could not reject the request (for details see [Difrancesco et al., 2012]). The time interval between "A makes a request" and "Successful communication" was considered an Interaction Time (Figure 2D).

The interaction was based on turns. Object selection for both request and response required the players to place the avatar's hand over the target object for three consecutive seconds. To facilitate the interaction, all objects lit up when they were being reached for. Both the selection of an object to be requested and the subsequent selection for delivery was indicated by a vertical rotation. Contrarily, when an incorrect response was provided (player B was reaching for an object different than the one requested), the chosen item lit up in red. All moves were continuously stored by the system.

To promote the activation of the language network including motor and associative cortices, the RGSa paradigm allowed using gestures and self-cueing strategies which were accompanied by but did not substitute, verbal communication [Cao et al., 1999, Pulvermüller and Fadiga, 2010]. Consequently, patients could and were encouraged to use both arms. Importantly, to counteract the use of compensatory strategies, the tables were separated by a compartment occluding the opponent's body movements [Difrancesco et al., 2012].

In the control group, SLT was provided by the public health-care system. Patients received an intensive standard SLT targeting individual linguistic deficits in an individual therapist-patient setting. Patients trained naming, repetition, spelling, and articulation of all target words. Both, intensity and frequency of the training in both groups were identical. Treatment was provided for 8 weeks, 5 days per week. The duration of each session was 30-40min, and the therapists ensured the same time of treatment for the two groups. Similarly, both groups used the same training materials: EG-122 3D objects (Figure 2B-C); CG-122 images presented in the form of cards. Additionally, the CG was encouraged to practice naming of the trained stimuli at home.

Statistical analysis

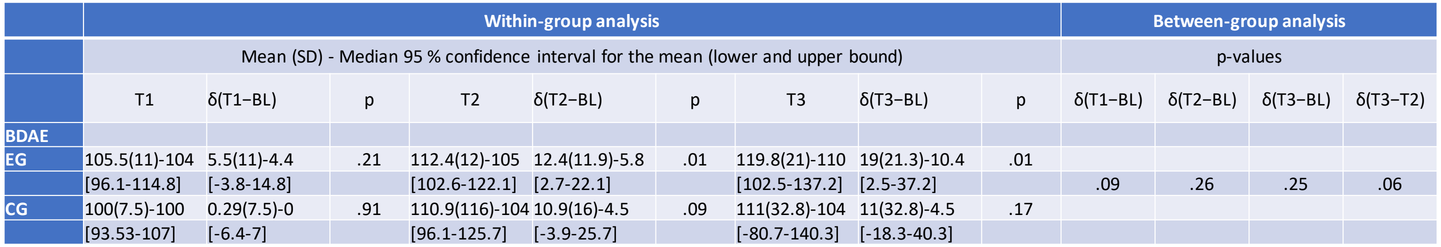
We used the Friedman test to compare the overall effect of treatments within the groups and the Wilcoxon signed-rank test to compare specific time steps. The Mann-Whitney U-test was performed to identify differences between groups. All comparative analyses used two-tailed tests and a standard level of significance ($p < .05$).

7.5.2 Supplemental Results

We assessed the homogeneity of the groups at T0 regarding lesion characteristics, demographics, and clinical variables (Supplement, Table2). The homogeneity of the groups was confirmed for all measures. Crucially, the analysis demonstrated no differences between EG and CG on BDAE ($p = .11$), CAL ($p = .36$), VocabT ($p = .29$), and FMA-UE (p=.14).

Although the statistical analysis yielded homogeneity of the groups at baseline for the BDAE ($p = .11$), it is notable that the scores were (non-significantly) higher in the EG (68.66) than CG (53.93). Therefore, to ensure a fair comparison between the groups on the primary outcome measure, we also compared the percentage change at each time step within and between groups (Supplement, Table4). The analysis yielded within-group changes for EG at the end of the treatment and the follow-up ($p = .21, p = .01, andp = .01$ respectively). We found neither statistically significant improvements for the CG ($T1 : p = .91, T2 : p = .09, T3 : p = .17$) nor differences between the groups ($T1 : p = .09, T2 : p = .26, T3 : p = .25$).

To determine the efficacy of the Interaction Times (ITs) measure as an objective quantification tool of improvement in lexical execution, we analyzed the linear relationship between ITs and performance on the VocabT. To perform the analysis, for each participant, we stored the total score on the VocabT at randomization, week 0, 2, 4, 6, and 8 together with the averaged IT extracted from the system at the corresponding time steps ± 1 day.



	Within-group analysis									Between-group analysis			
	Mean (SD) - Median 95 % confidence interval for the mean (lower and upper bound)									p-values			
	T1	δ(T1−BL)	p	T2	δ(T2−BL)	p	T3	δ(T3−BL)	p	δ(T1−BL)	δ(T2−BL)	δ(T3−BL)	δ(T3−T2)
BDAE													
EG	105.5(11)-104	5.5(11)-4.4	.21	112.4(12)-105	12.4(11.9)-5.8	.01	119.8(21)-110	19(21.3)-10.4	.01				
	[96.1-114.8]	[-3.8-14.8]		[102.6-122.1]	[2.7-22.1]		[102.5-137.2]	[2.5-37.2]		.09	.26	.25	.06
CG	100(7.5)-100	0.29(7.5)-0	.91	110.9(116)-104	10.9(16)-4.5	.09	111(32.8)-104	11(32.8)-4.5	.17				
	[93.53-107]	[-6.4-7]		[96.1-125.7]	[-3.9-25.7]		[-80.7-140.3]	[-18.3-40.3]					

Figure 7.7: Percentage change on BDAE at month 1 (T1), at the end of treatment (T2), and follow-up (T3). Bold values indicate significant differences ($p < .05$). P-values for within-group analysis were obtained with Wilcoxon signed-rank test while p-values for the between-group analysis were obtained with Wilcoxon rank-sum respectively.

7.5.3 Supplemental Discussion and Limitations

An extensive body of research supports the effectiveness of SLT even in the chronic phases of aphasia and sheds light on variables driving neurological and functional recovery [Berthier and Pulvermüller, 2011, Stahl et al., 2016, Katz et al., 2000, Pulvermüller et al., 2016]. In particular, studies support the positive impact of four major therapeutic components: (1) high training intensity and frequency, (2) language training in the context of behaviorally relevant and goal-oriented speech-acts, (3) avoidance of functional compensation by preventing learned non-use, and (4) social-communicative embedding of language training [Berthier and Pulvermüller, 2011, Stahl et al., 2016, Pulvermüller et al., 2016, Pulvermüller and Berthier, 2008]. From a clinical perspective, The World Health Organization (WHO) proposes that a complete rehabilitation program should consist of Disorder-Oriented Treatment (DOT), Functional Treatment (FT) and Participation-Oriented Treatment (POT) [mondiale de la santé and Organization, 2001] in either standard or computer-based manner, however, such method is rarely practiced by social health-care [Thompson et al., 2008]. To provide lexical and syntactic training in a multimodal, goal-oriented way within a context of dyadic peer interaction, RGSa was designed to integrate these domains. In particular, our paradigm had a form of a language game involving peers without a continuous active involvement of a therapist. Furthermore, exposure to the multisensory cues, both visuomotor and auditory, within the action-context aimed to trigger the activation of the language network targeting use-dependent neuroplasticity [Pulvermüller and Fadiga, 2010, Stahl et al., 2016, Pulvermüller and Berthier, 2008, Kleim and Jones, 2008]. At the current stage, we cannot disambiguate between the contribution of specific variables of the RGSa to the reported changes to determine their impact. However, similar to others [Stahl et al., 2016] we hypothesize that patients in the RGSa group might have particularly benefitted from the dyadic nature of the therapy embedded within behaviorally-relevant context. In particular, the peer interaction of the RGSa, which required the use of language in everyday-life situations, could have positively influenced the frequency of language use in social activities outside of the hospital thus reinforcing the trained vocabulary and preventing non-use. Although at this stage the specific mechanisms through which such peer-peer therapeutic context could have influenced the reported outcomes remain elusive, they are likely multifactorial. First, it is noteworthy that animal research suggests a significant role of social interaction for neural and functional recovery post-stroke [Craft et al., 2005, Venna et al., 2014]. Second, motor control literature proposes that greater functional outcomes and positive neuroplastic changes are more likely when the training incorporates complex, as compared to simple, tasks and environments [Komitova et al., 2005, Ding et al., 2003]. Hence, outcomes of the RGSa treatment might have also been affected by the dynamics of the Language-Action Game (LAG) as well as the enhanced VR-based environment. Finally, patients themselves could have provided mutual support and encouragement which was effective because they could understand each other’s dilemmas and frustrations related to the communication after stroke. In that case, the interaction in itself could be therapeutic and could contribute to the effectiveness of the reported results. On the other hand, a recent RCT which compared ILAT with an intensive naming therapy (both in a group setting) allowing social interaction and communication with other patients, showed that ILAT led to a significantly better language outcome than the control group [Pulvermüller and Fadiga, 2010]. This result might suggest that, while social interactions can have a positive influence on language performance, the beneficial effects of ILAT cannot be explained purely by this factor. Future studies should systematically investigate the contribution of the discussed variables including multisensory cueing (i.e., visual, auditory, haptic) on functional and structural recovery as well as their both short and long-term effects. Moreover, comparing the RGSa method with the original ILAT would shed light on the effects of the computer-based nature of the intervention.

To assess whether at the current stage the proposed RGSa therapy can already be employed to the clinic without assistance, all the sessions were supervised by a Therapy Assistant (TA). Their primary role was to monitor possible technological or communication difficulties and report them to the experimenters. Additionally, the TA aided the patients when a trial could not be completed independently. However, they did not provide any additional cues. TA reported no issues related to either the technical aspects of the system or the interaction between the participants. This feedback suggests that the system is robust and safe to be integrated within a clinical setting to complement the conventional treatment and increase its intensity without the guidance of a TA. Within such context, future studies shall investigate the effects of a combined therapy which integrates standard SLT and the RGSa training. Such rehabilitation protocol might be highly beneficial for patients with aphasia targeting individual linguistic needs on the one hand and behaviorally relevant peer-interaction on the other. Additionally, one of the objectives of virtual-reality based therapy, such as the RGSa, would be the use as a supplemental training device for patients in the clinic or at home. In fact, in most industrialized countries, healthcare resources for chronic patients with aphasia are very limited, although ample research shows language improvements even at the chronic stage of the disease. Therefore, in the future, RGSa may have the potential to help those patients who are no longer offered any therapeutic support by the public health care system.

The computer-based nature of the RGSa has an additional advantage of further testing its plausibility and effectiveness at home. Specifically, the significant relationship between the ITs and VocabT strongly supports the notion that ITs reflect the recovery of lexical access and verbal execution. First, such design might promote language use and aid in overcoming social isolation after hospital discharge. Second, the interaction through the internet and remote access to patients' data including ITs could provide continuous information on the recovery and allow tailoring of the trained stimuli to the individual needs of the patients. Crucially, to provide a complete linguistic assessment, a future version of the proposed system should implement speech recognition tools allowing to determine specific pronunciation errors [Witt, 1999].

In conclusion, we have shown that RGSa, based on principles of use-dependent learning, behavioral relevance, and treatment intensity leads to functional linguistic recovery that is comparable to standard SLT at the end of the intervention and superior in the long-term follow-up. For these reasons, and in line with others [Palmer et al., 2013, Palmer et al., 2012, Stark and Warburton, 2016, Katz and Wertz, 1997, Kurland et al., 2018, van de Sandt-Koenderman, 2011], we believe that the current challenges faced by the health-care system in the treatment of stroke can be effectively addressed by augmenting traditional therapy with computer-based methods.

7.5.4 Full List of the Therapy Items

Spanish: Abeja, acordeón, aguacate, ajo, alcachofas, amapola, ambulancia, armónica, arpa, aspiradora, avión, banco, banjo, barco, bateréa, batidora, bellota, bocadillo, bomba, brújula, bus, cactus, calabaza, calcetín, cama, camiseta, camión, campana, caracol, casa, calcetín, cepillo de dientes, cereales, chaqueta, coche, cohete, comida, copa, cortacésped, cremallera, cubeta, donuts, ducha, embudo, escalera, flauta, fresas, fuego, gallo, gancho, gato, gorro, grabadora, grifo, guisantes, guitarra, helicóptero, hoja, iglú, jarra, jeringa, jersey, lazo, leche, limonada, llaves, locomotora, maquinilla de afeitar, maracas, mariposa, melocotón, mesa, mora, mosca, moto, máquina de coser, máquina de escribir, paleta de pintura, palmera, pantalones, papaya, pastel, peine, pelota, pera, perro, piano, pimiento, pinball, pipa, pirámide, piscina, pistola, piano, plátano, prismáticos, radio, regalo, reloj, sandía, saxofón, seta, silbato, sillón, taladro, tanque, teclado, teléfono, tocadiscos, tomates, tractor, tranvía, tren, trompeta, tv, uvas, vaca, ventilador, wáter, órgano.

English: Bee, accordion, avocado, garlic, artichokes, poppy, ambulance, harmonica, harp, vacuum cleaner, plane, bench, banjo, boat, battery, blender, acorn, sandwich, bomb, compass, bus, cactus, pumpkin, sock, bed, t-shirt, truck, bell, snail, house, sock, toothbrush, cereals, jacket, car, rocket, food, cup, lawnmower, rack, bucket, donuts, shower, funnel, ladder, flute, strawberries, fire, rooster, hook, cat, cap, burner, faucet, peas, guitar, helicopter, leaf, igloo, jug, syringe, jersey, lasso, milk, lemonade, keys, locomotive, razor, maracas, butterfly, peach, table, blackberry, fly, motorbike, sewing machine, typewriter, paint palette, palm tree, pants, papaya, cake, comb, ball, pear, dog, piano, pepper, pinball, pipe, pyramid, pool, gun, piano, banana, binoculars, radio, gift, watch, watermelon, saxophone, mushroom, whistle, armchair, drill, tank, keyboard, telephone, turntables, tomatoes, tractor, tram, train, trumpet, tv, grapes, cow, fan, toilet, organ.

7.5.5 Full List of Items from the CAL

Spanish:
1. Con que frecuencia el paciente se comunicaba con un familiar o un buen amigo?
2. Con que frecuencia se comunicaba cuando estaba en un grupo de familiares o amigos?
3. Con que frecuencia se comunicaba con un extranjero?
4. Con que frecuencia se comunicaba cuando estaba en un grupo con alguna gente que no conocía?
5. Con que frecuencia se comunicaba en una oficina, tienda, o institución publica (postal etc.)?
6. Con que frecuencia usaba el teléfono?
7. Con que frecuencia escuchaba las noticias en la radio o en la televisión?
8. Con que frecuencia leía el periódico/revista?
9. Con que frecuencia hacía apuntes?
10. Con que frecuencia hacía simples problemas aritméticos?
11. Con que frecuencia se comunicaba cuando estaba bajo estrés?
12. Con que frecuencia se comunicaba cuando estaba relajado, o no estaba bajo estrés?
13. Con que frecuencia se comunicaba cuando estaba cansado/a?
14. Con que frecuencia decía su opinión o reportaba hechos?
15. Con que frecuencia hacía preguntas?
16. Con que frecuencia respondía a las preguntas hechas por los demás?
17. Con que frecuencia verbalmente expresaba críticas, o lamentaba?
18. Con que frecuencia verbalmente respondía a críticas?

English:
1. How frequently would the patient communicate with a relative or good friend?

2. How frequently would the patient communicate when together with a group of friends or relatives?
3. How frequently would the patient communicate with a foreigner?
4. How frequently would the patient communicate when in a group together with several others he or she does not know?
5. How frequently would the patient communicate in an office, store or public institution (post office, butcher etc.)?
6. How frequently would the patient use the telephone?
7. How frequently would the patient listen to news on the radio or TV?
8. How frequently would the patient read the newspaper?
9. How frequently would the patient write down short notes?
10. How frequently would the patient solve simple arithmetic problems?
11. How frequently would the patient communicate when under stress?
12. How frequently would the patient communicate when relaxed and not under stress?
13. How frequently would the patient communicate when he or she is tired?
14. How frequently would the patient make statements or reports about facts?
15. How frequently would the patient ask a question?
16. How frequently would the patient answer questions asked by others?
17. How frequently would the patient verbally express criticisms or make complaints?
18. How frequently would the patient verbally respond to criticisms?

Chapter 8

THE EFFECTS OF SILENT VISUOMOTOR CUEING ON WORD RETRIEVAL IN BROCA'S APHASICS: A PILOT STUDY

This chapter is based on:
Grechuta, K., Ballester, B. R., Munné, R. E., Bernal, T. U., Hervás, B. M., San Segundo, R., & Verschure, P. F. (2017, July). The effects of silent visuomotor cueing on word retrieval in Broca's aphasies: A pilot study. *In Rehabilitation Robotics (ICORR), 2017 International Conference on (pp. 193-199).* IEEE.

About a quarter of stroke patients worldwide suffer serious language disorders such as aphasias. Most common symptoms of Broca's aphasia are word naming disorders which highly impact verbal communication and the quality of life of aphasic patients. In order to recover disturbances in word retrieval, several cueing methods (i.e. phonemic and semantic) have been established to improve lexical access establishing effective language

rehabilitation techniques. Based on recent evidence from action-perception theories, which postulate that neural circuits for speech perception and articulation are tightly coupled, in the present work, we propose and investigate an alternative type of cueing using silent articulation-related visual stimuli. We hypothesize that providing patients with primes in the form of silent videos showing lip motions representative of correct pronunciation of target words, will result in faster word retrieval than when no such cue is provided. To test our prediction, we realize a longitudinal clinical virtual reality-based trial with four post-stroke Broca's patients and compare the interaction times between the two conditions over the eight weeks of the therapy. Our results suggest that silent visuomotor cues indeed facilitate word retrieval and verbal execution, and might be beneficial in lexical relearning in chronic Broca's patients.

8.1 Introduction

Stroke often results in speech impairments which highly impact patients' quality of life. Due to natural recovery processes and treatment, the brain might repair both at early and chronic stages after the lesion onset [Goodglass, 1993, Xerri et al., 1998]. Indeed, a number of stroke survivors regain lost functions, which allows them to succeed in the activities of daily living (ADLs) [Mayo et al., 2002]. The majority of patients, however, are left with long-term disabilities leading to dependence, isolation and depression [Whyte and Mulsant, 2002], which can adversely affect recovery [Hadidi et al., 2009].

Aphasia, in particular, is a disturbance of language processing including comprehension, production and memory, caused by a neurological injury, most commonly, stroke [Lichteim, 1885]. Here, we support that aphasia, as any other functional disturbance resulting from a neurological injury, is a brain disease. Consequently, the treatment of such a deficit should comprise knowledge from both behavioral and neurophysiological literature [Berthier and Pulvermüller, 2011], as well as our computational understanding of the brain [Verschure, 2011].

One of the most common symptoms of Broca’s aphasia is difficulty in word retrieval and naming (verbal execution). In order to overcome subsequent disturbances, a number of cueing methods (i.e. semantic and phonetic) were established, tested and applied to improve both immediate and long term lexical access [Howard and Papathanasiou, 2000]. During treatment, both semantic and phonemic cues are usually administered by the therapist in a written or oral form. They act as primes containing phonological, semantic or syntactic information relevant to target lexicon. Phonemic cues indicate the initial sound/s of target words, while semantic cues provide description of target words. It was suggested that, among all, Broca’s and conduction aphasia patients show the highest responsiveness to cueing [Lowell et al., 1995]. Indeed, both phonemic and semantic priming were shown highly beneficial in word retrieval trainings and are widely practiced in clinical setups [Abel et al., 2007]. In the following sections, we discuss evidence from neurophysiological research [Petrides et al., 2005, Nishitani and Hari, 2002] which leads us to propose and test an alternative type of cueing, namely, silent visuomotor (SVC).

First, it was demonstrated that motor circuits which were traditionally thought to subserve motor planning and control seem to be also involved in the comprehension and perception of language [Watkins and Paus, 2004, Pulvermüller et al., 2006]. In particular, processing of action words (i.e. related to hand, leg, or lip movements) was shown to evoke immediate somatotopic activation of the motor cortex [Hauk et al., 2004]. The structures include populations of neurons which are involved in action execution, observation, and imagination. On the other hand, results from recent studies in human and nonhuman primates show that Broca’s area, which was originally thought to play a major, if not exclusive, role in language production [Lichteim, 1885, Damasio and Geschwind, 1984] is also involved in other language-related functions [Fadiga et al., 1995, Fadiga et al., 2002, Pulvermüller et al., 2003, Arbib, 2010].

On one hand, neurophysiological evidence suggests the existence of mirror neurons in Broca’s area and their activation during the recognition of hand gestures [Fadiga et al., 1995, Rizzolatti and Arbib, 1998], processing of articulatory phoneme-specific muscles during speech perception [Fadiga et al., 2002] as well as language comprehension [Pulvermüller et al., 2003]. Other human studies have further shown the involvement of Broca's area in observation [Rizzolatti et al., 1996], preparation [Krams et al., 1998], and imagination [Grafton et al., 1996] of goal-oriented movements which involve hand. These results imply that Broca's region can subserve sensorimotor matching during not only language but also motor processing [Arbib, 2010], which gave rise to novel approaches to treatment of language disorders such as the Intensive Language-Action Therapy [Difrancesco et al., 2012].

Interestingly, it was recently proposed that the area F5 in the macaque monkey is neuroanatomically comparable to the Brodmann area 44 (BA 44), which forms part of Broca and, in macaques, this homologue of human Broca's area responds to orofacial gestures [Petrides et al., 2005]. Similarly as in the macaque monkey, in humans, area 44 lies just in front of the agranular premotor cortex which controls orofacial musculature. The specific question of whether Broca's area is sensitive to silent, still (i.e. motionless) orofacial stimuli in humans was further tested with healthy subjects using fMRI [Nishitani and Hari, 2002]. The authors compared the activation of human mirror-neuron system (MNS), including Broca's area, during observation, imitation and execution of verbal and nonverbal lip forms. The results show that the signals from motor cortex and Broca's area were significantly stronger during imitation than other conditions suggesting involvement of the area in the processing of orofacial gestures. Following these findings, in the present study, we investigate whether silent visuomotor cueing has a beneficial effect on word retrieval and verbal execution of target words, and whether it influences relearning of lexicon in Broca's aphasics.

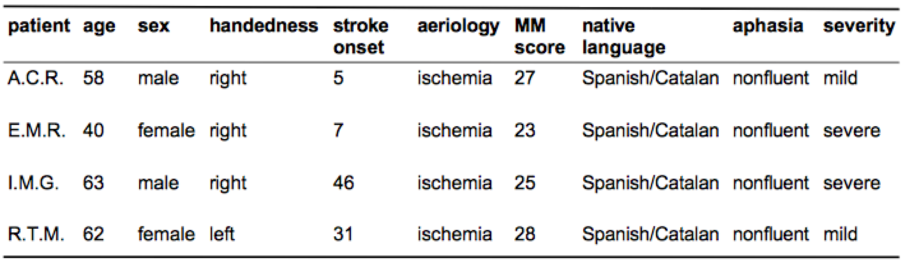
Thanks to technological advancements a number of scientifically grounded computer-assisted treatments have emerged as novel language rehabilitation paradigms [Fink et al., 2005]. Recent evidence suggests that both simple phonemic cueing [Bruce and Howard, 1987] and complex multi-cueing including semantic, phonologic, and orthographic hints [Kurland et al., 2014], generated by a computer, are promising methods for treating word retrieval disorders in aphasias. This supports that computer-assisted language therapies can be used with patients either to intensify a conventional therapy, or to reinforce the treatment at home. Here, we implement silent visuomotor cues (SVCs) in a form of videos representing lip motions representative for correct pronunciation of target words within a virtual reality (VR)-based rehabilitation environment. We test four chronic post-stroke Broca's patients in a longitudinal clinical trial. We hypothesize that exposure to SVCs will facilitate word retrieval and verbal execution of the target words. To test our prediction we compare the time it takes for each patient to successfully request the target object from an opponent and receive it (i.e. interaction times (IT)). We expect that the IT's will decrease in both experimental (cued stimuli) and the control conditions (non-cued stimuli), over the therapy interval. Nevertheless, we predict that overall the ITs will be lower in the cued trials, specially in the early intervention days when the exposure to the lexicon is still low.

8.2 Methods

8.2.1 Subjects

Four chronic (Mean(SD): 22.25±19.75 months post-stroke) Broca's aphasia patients (Figure 8.1) participated in the study (A.C.R., E.M.R., I.M.G., R.T.M.; age: Mean(SD): 55.75±10.71). All the participants had normal or corrected-to-normal vision. We excluded patients with major perceptual, motor and neuropsychological impairments such as severe motor dysfunction, apraxia, visual deficits, neglect, executive dysfunctioning, dementia or attentional disorders. Representative MRI images of patients' lesions are shown in Figure 8.2. The images demonstrate significant damage throughout the left hemisphere including Broca's area.

The reported experimental procedures followed written consents from all the involved patients and had been officially registered and accepted. The study was further approved by the local Ethical Committee from the Hospital Universitari Joan XXIII.



patient	age	sex	handedness	stroke onset	aeriology	MM score	native language	aphasia	severity
A.C.R.	58	male	right	5	ischemia	27	Spanish/Catalan	nonfluent	mild
E.M.R.	40	female	right	7	ischemia	23	Spanish/Catalan	nonfluent	severe
I.M.G.	63	male	right	46	ischemia	25	Spanish/Catalan	nonfluent	severe
R.T.M.	62	female	left	31	ischemia	28	Spanish/Catalan	nonfluent	mild

Figure 8.1: Sociodemographic and clinical data of A.C.R., E.M.R., I.M.G., R.T.M.. MM: The Mini-Mental State Examination (max= 30 points).

8.2.2 Experimental Setup

The experimental setup (Figure 8.5 A) consisted in 2 computers connected via a local network and 2 motion sensors (Kinect, Microsoft, Seattle). In a therapy ward, each patient was seated at a cut-out table, facing the screen. To prevent the users from seeing each others’ displays and hand movements, the tables were placed in front of each other. Patients interacted with the system by performing horizontal arm movements over the table surface. These movements were tracked in real time and mapped onto avatars’ upper-limbs, allowing the interaction with virtual objects. All the patients were provided with a wireless headset through which they heard feedback from the game (i.e. change of turns or correct pronunciation of target words). Detailed, technical description of the setup can be found in [Grechuta et al., 2016a].

8.2.3 Experimental Procedure

All the patients went through the same therapeutic paradigm following a within-subjects design. The proposed therapy is a multiplayer language game (i.e. Sensorimotor Speech Therapy, SST), integrated within a VR environment for motor and cognitive training, the so-called Rehabilitation Gaming System (RGS) [Cameirão et al., 2010,Grechuta et al., 2016a]. The paradigm involves two patients interacting with each other by requesting objects. To prevent from the potential novelty effects, prior to the intervention, all the patients are familiarized with the system. During the first

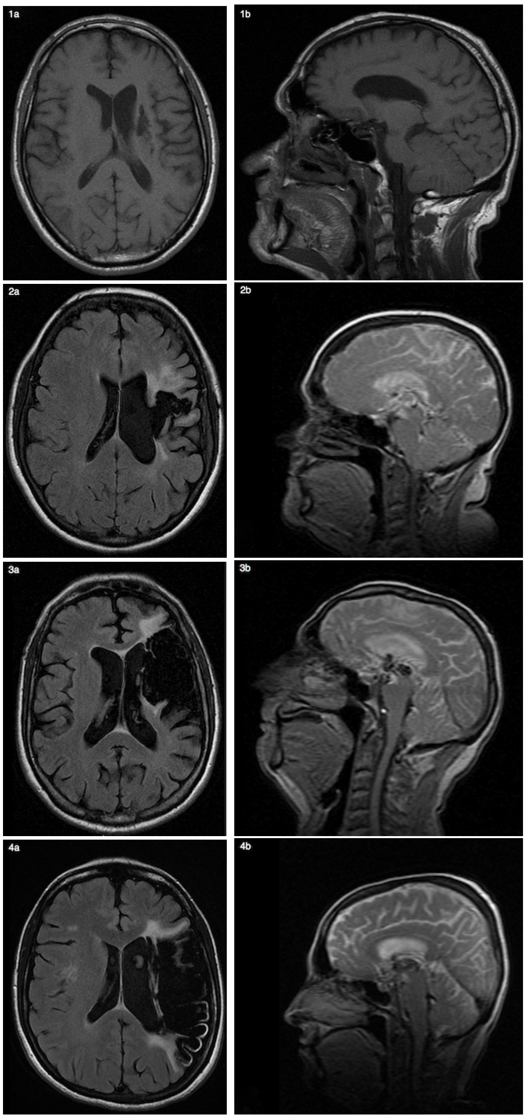


Figure 8.2: MRIs. Left: images of axial brain slice with maximum ischemic damage for each patient. Right: sagittal slices through left hemisphere. Images 1a and 1b: patient A.C.R. (mild aphasia); 2a and 2b: R.T.M. (mild aphasia); 3a and 3b: I.M.G. (severe aphasia); 4a and 4b: E.M.R. (severe aphasia).

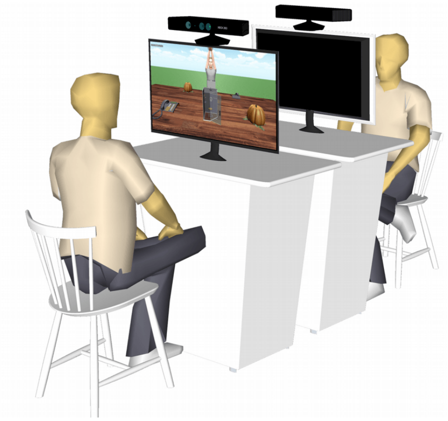


Figure 8.3: Therapeutic setting. Two patients sit on a chair, with their arms resting on a table and facing the computer screen. The movements of the arms are continuously captured by a vision-based motion tracking system (Kinect, Microsoft, Seattle). Tracked movements are mapped onto the movements of an avatar, mimicking the movements of the user and enabling interaction with the objects.

Figure 8.4: 3D stimuli. The objects depict different stimuli categories: (1) frequency of lexicon including (a) high frequency ("bed", "house"), (b) middle frequency ("milk", "lunch") and (c) low frequency ("cactus", "pumpkin"); (2) objects from the same semantic categories (i.e. food); (3) simple and complex objects which require more than two words (simple: "garlic", Spanish "ajo" and complex: "toothbrush", Spanish "cepillo de dientes"); (4) objects which differ by one phoneme ("bed" and "house", Spanish "casa" and "cama").

Figure 8.5: Experimental setup (a) and therapy materials (b).

session, the patients are given detailed instructions about the experimental procedure. One of the patients is randomly assigned to begin the game. The goal of the game for each patient and each session is to request and collect objects from the opponent, or give them, when requested. Each trial is divided into 3 steps: 1) the player selects the desired object by placing the hand on top of it, 2) s/he verbally requests it and 3) the opponent understands the request and gives the corresponding object by placing their hand on top of it. After a successful trial, the patient receives points, hears the correct pronunciation of the target word and the turn changes. There is a set of three objects per patient simultaneously available for selection and both sets are always visible for both players (Figure 8.5 A). During every session each patient is to request and collect 36 objects. All the manipulations within the game (i.e. object selected, selection time, giving time, etc.) are continuously recorded and stored by the system for further analysis.

To test our hypothesis and determine whether SVCs facilitate word retrieval, word articulation, and have beneficial effects on lexical relearning, 50 % of the stimuli (n=30) was provided with silent videos representing the correct pronunciation of target words. The representative videos appeared on the screen immediately after object selection. The cues were presented only once per trial.

Here, we report preliminary results form an ongoing longitudinal clinical trial taking place in the Clínica de l'Hospital Universatari Joan XXIII, Tarragona, Spain. The experimental intervention lasted 8 weeks and consisted in 5 daily training sessions with the present system, each of them with a duration of approximately 40 minutes.

8.2.4 Stimuli

The materials consisted in 60 3D objects adapted from [Snodgrass and Vanderwart, 1980]. Examples of the stimuli are shown in Figure 8.5 B. In order to ensure that the stimuli were visually unambiguous and likely to elicit the desired target words, they were first evaluated by healthy subjects and the language therapists who supervised the treatment. The selected stimuli were emotionally neutral and were categorized according to frequency (i.e. high, middle and low frequent occurrence in Spanish language), semantic category, complexity and phonemic similarity (i.e. minimal pairs). Lexicon was further matched for syllable length and semantic field. For each pair of patients, the therapy items were randomly split into two sets, so that half of the stimuli was accompanied with the visuomotor cue and half was not. The stimuli were presented in a pseudo-randomized order, counterbalanced within each week.

8.2.5 Cues

The videos were recorded in the Clínica de l'Hospital Universitari Joan XXIII de Tarragona, Spain. We recorded a language therapist who read out-loud each of the 60 stimuli. The reading followed the criteria of standard phonetic cueing. However, instead of the first phoneme/s only, the therapist pronounced full target words. For the purpose of the present experiment, we muted the voice so that the recordings were silent. The videos depicted part of the face of the therapist including mouth and nose 8.6.

Figure 8.6: Image representing an example of the stimuli.

8.2.6 Measures

To evaluate the effects of SVCs on word retrieval and verbal execution, we extracted and computed the averaged Interaction Times (ITs) for all the stimuli and for all the therapy sessions. We defined the IT as the time interval between the selection and collection of an object; both automatically stored by the system. Each IT included word retrieval, articulation of the request and comprehension of the target word by the opponent.

To estimate the learning of the target stimuli, we administered a vocabulary test (VocabT) prior to the intervention and at end of every second week of the intervention (N=5). VocabT was performed by a blinded therapist and included all the 60 items. For each word patients could score a maximum of 5 points (0p: no verbal utterance, 1p: utterance followed by full phonetic priming, 2p: utterance followed by priming of initial phoneme, 3p: utterance followed by full silent orofacial hint, 4p = utterance followed by silent orofacial hint of the first phoneme, 5p: utterance followed by no hint). The maximum score for each test was 300 points.

8.3 Results

The goal of this study was to determine the effects of SVCs on verbal execution in chronic Broca’s patients. To test the overall effectiveness of the intervention we first computed the mean ITs from all sessions per each patient including cued and non-cued stimuli, and correlated the results with the performance on the VocabT. Subsequently, to estimate whether the proposed priming facilitates verbal expression, we compared the ITs in cued and noncued stimuli groups for all the sessions as well as for the early and late trials. Here, we excluded first 3 sessions to account for possible novelty effects. Normality test showed that the distributions of ITs for all the patients were not normal. Consequently, the statistical analysis followed nonparametric analysis.

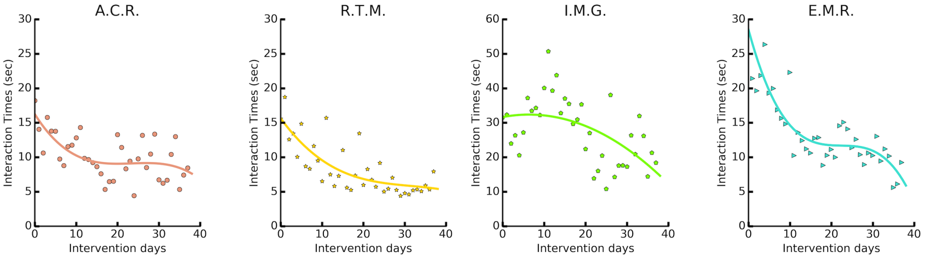


Figure 8.7: Mean Interaction Time. Mean IT for all trials (cued and non-cued) per session, over the therapy interval.

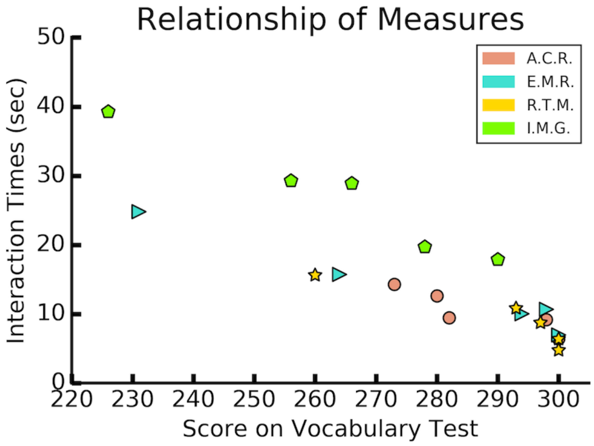


Figure 8.8: Measures’ correlation. X-axis: VocabT; y-axis: ITs.

Our data shows that in all the trials (both cued and non-cued) the ITs decreased for all the patients over the therapy interval (Figure 8.7). A Pearson product-moment correlation was run to determine the linear relationship between ITs and the performance on the VocabT for each patient (Figure 8.8). The analysis revealed strong, statistically significant, negative, correlation between ITs and VocabT scores for A.C.R ($rs = -.88$, $p < .05$), R.T.M. ($rs = -.91, p < .05$), I.M.G. ($rs = -.97, p < .005$) and E.M.R. ($rs = -.98, p < .005$).

A Wilcoxon signed-ranks test was performed to compare the overall ITs between the experimental (cued) and the control (non-cued) trials (Figure 8.9). The results demonstrate statistically significant differences between the two conditions in case of patients with severe aphasia: I.M.G ($Z = 167, p = .01$) and E.M.R. ($Z = 84, p > .001$). No such differences were found for neither A.C.R. ($Z = 289, p = .67$) nor R.T.M. ($Z = 277$, $p = .53$).

To further estimate the effects of SVCs on verbal execution, we computed mean ITs for the early and the late trials in the experimental and control conditions for every patient 8.10. A Wilcoxon signed-ranks test revealed significant differences between early and late trials in non-cued stimuli in both A.C.R. ($Z = 8, p > .04$) and R.T.M. ($Z = 1, p > .006$).

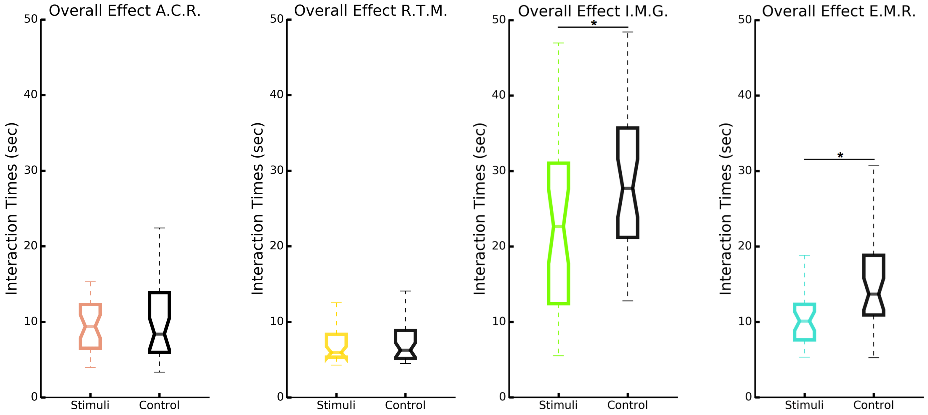


Figure 8.9: Overall Mean Interaction Times. Box plots represent mean ITs for all the therapy (N=38 days) for each individual patient for cued (colored) and non-cued (black) trials.

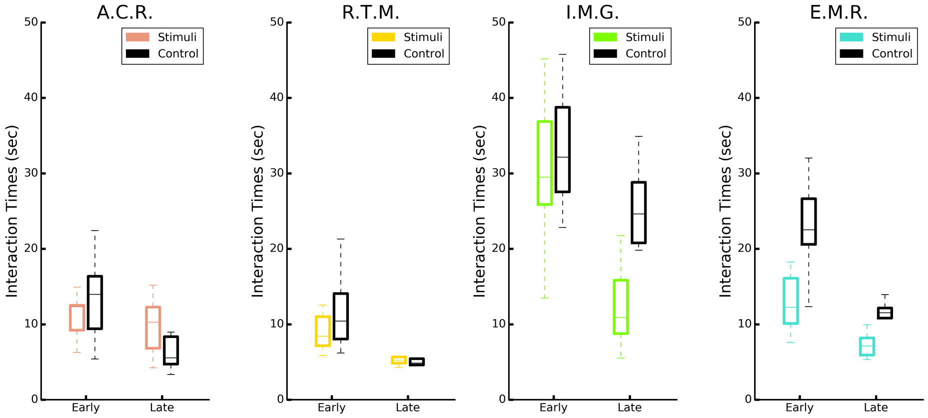


Figure 8.10: Mean Interaction Times. Mean IT for each individual patient in early (days 3:13, N=10) and late (28:38, N=10) sessions. Color indicates cued, and black non-cued trials.

Such difference was also found for the cued trials in case of R.T.M ($Z = 2$, $p > .009$), but not A.C.R.. No differences were found within early and late trials between cued and non-cued stimuli for neither of the patients with mild aphasia (A.C.R. and R.T.M.).

In case of E.M.R. (severe aphasia) we report significant difference in ITs between cued and con-cued stimuli in the early sessions (E.M.R.:$Z = 0$, $p > .005$). We have not found such difference in case of I.M.G. (I.M.G.: $Z = 23, p > .64$). We further show significant differenced between cued and non-cued stimuli in the late sessions for both severe patients (I.M.G.: $Z = 1, p > .006$; E.M.R.:$Z = 3, p > .01$). Finally, Wilcoxon signed-ranks test revealed differences between early and late trials for cued (I.M.G.: $Z = 2, p > .009$; E.M.R.:$Z = 0, p > .005$) and non-cued stimuli for both patients (I.M.G.: $Z = 8, p > .04$; E.M.R.:$Z = 0, p > .005$).

8.4 Discussion

Stroke can provoke language disorders which affect patients' quality of life. Broca's patients, in particular, often suffer nonfluent speech and difficulties in verbal execution. The goal of clinical rehabilitation of those deficits is to retrieve lost functions and prevent learned nonuse [Meinzer et al., 2007, Ballester et al., 2015a] through supervised training. Only recently, however, standard clinical methods started to be tested against novel hypotheses from neurophysiological research in primates [Davis and Johnsrude, 2003]. Interestingly, several findings shed new light onto the nature of neural networks which govern language processing, both perception and production [Graziano et al., 2002, Nishitani and Hari, 2002, Petrides et al., 2005, Pulvermüller et al., 2006, Fadiga et al., 2009]. Here, we propose that evaluating biologically-based premises in behavioral paradigms can be beneficial for post-stroke recovery providing further insights on methods that drive language relearning and improving current rehabilitation paradigms. Indeed, a number of protocols following this assumption have recently shown to successfully promote repair of neural circuits enhancing the quality of life of aphasics [Naeser et al., 2005, Richards et al., 2008, Pulvermüller and Berthier, 2008, Edmonds et al., 2009].

Here, we followed this line of research and behaviorally tested recent evidence from action-perception theories which postulate that neural circuits for speech perception and articulation are coupled [D'Ausilio et al., 2009]. In particular, we investigated the effects of SVC on word retrieval in chronic Broca's aphasia patients, two severe and two mild. As expected, the ITs decreased for all the stimuli during the intervention in all patients (Figure 8.7), which suggests an overall efficacy of the proposed, longitudinal treatment. It is worth to note, however, that the effect was stronger for the severe (A.C.R. and R.T.M.) than for the mild patients (A.C.R. and R.T.M.). The decrease of the ITs was significantly and highly correlated with the increase in performance on the VocabT in all patients. We defined IT as time interval between the selection and collection of objects, which included word retrieval, verbal execution, comprehension and related motor actions. Thus we proposed ITs as an objective measure of interaction (i.e. obtaining a goal), rather than linguistic performance. The reported correlation with the clinical scale, however, suggests that the ITs could be regarded as an unsupervised measure of improvement in verbal execution. More patients are being tested to further validate this hypothesis.

In line with our hypothesis was the difference in ITs between cued and non-cued trials. Interestingly, however, the effect was statistically significant in case of severe patients I.M.G. and E.M.R., while in case of A.C.R. and R.T.M. we report a trend. Similarly, the ITs between cued and non-cued stimuli in late trials were significantly different in case of severe patients. These outcomes might be influenced by a ceiling effect in patients with mild aphasia. Finally, we observe a trend in the early trials between the cued and control stimuli suggesting a beneficial effects of the stimuli. This result was statistically significant for E.M.R..

Present preliminary outcomes suggest that SVCs can be beneficial in severe cases of Broca's aphasia. Only one patients (A.C.R.) seemed not to have significantly improved in the cued condition. We suspect that this result might be affected by the severity of aphasia, stroke onset and high performance of the first VocabT, suggesting a ceiling effect. The remaining three patients E.M.R., I.M.G. and R.T.M. improved in the cued condition and benefited from cues in the early trials.

Previous evidence demonstrated that multimodal priming (i.e. auditory and visual) facilitates word retrieval in Broca's aphasics [Hall et al., 2005, Kendall et al., 2008, Lee et al., 2010, Fridriksson et al., 2012]. This method, also known as 'speech entrainment', enables patients to produce fluent speech in real time [Fridriksson et al., 2012]. Indeed, audiovisual priming showed significantly lower latency scores in word production task compared to unimodal, audio-only feedback or non-primed spontaneous speech [Fridriksson et al., 2012]. Results from multisensory speech perception studies highly encourage multimodal, both standard and VR-based, priming methods in treatment of Broca's aphasia. It was suggested, however, that audiovisual speech perception and silent speech-reading possibly share common processes of visual analysis [Hall et al., 2005]. Interestingly, Nishitani and colleagues [Nishitani and Hari, 2002] found that Broca's area is involved in the processing of silent orofacial gestures in healthy subjects. One of the critical aspects of the present protocol was that the stimuli were mute. Thus our behavioral results are in line with previous studies [Petrides et al., 2005, Nishitani and Hari, 2002] further supporting the hypothesis that language perception might consist in, not purely audiovisual [Reisberg et al., 1987], but also visuomotor processing and that Broca's area plays a major role in governing these processes [Nishitani et al., 2005, Skipper et al., 2007, Corballis, 2009, Arbib, 2010]. As such, we further propose that Broca's area is also a "perceptual system", as predicted in DAC5 [Verschure and Althaus, 2003] and [Wyss et al., 2004], which subserves perceptuomotor processing [Pulvermüller and Fadiga, 2010, Kohler et al., 2002].

Overall, our findings provide clinical evidence for beneficial effects of SVCs on verbal execution, which could extend current rehabilitation practice. Additionally, the proposed protocol can find applications in computer-based language rehabilitation systems providing an on-line assessment of performance and adjusting intensity and difficulty of training for personalized practice.

8.5 Conclusions

We proposed and validated an alternative priming strategy for chronic Broca's aphasics: silent visuomotor cueing. The experiment was integrated in a longitudinal clinical trial. Four patients underwent an eight-week VR-based sensorimotor language intervention in a form of a multiplayer game. To accomplish every session patients were to request and give objects. Prior to the object request, at the moment of selection, in half of the stimuli, patients were exposed to videos showing lip motions representative for correct pronunciation of target words. The protocol allowed for automatic, continuous quantification of the priming effects within the VR-based therapy. Our results show that the ITs, measured as time intervals between selection and collection of objects, were significantly lower in cued trials in patients with severe aphasia. First, this suggests that integrating SVC in the rehabilitation of severe Broca's aphasics might be beneficial for training word retrieval and verbal execution possibly triggering the activation of the language-motor brain areas related to perception of orofacial gestures. Secondly, the present outcome provides evidence for the multimodal integration at the core of language processes, which is consistent with that observed in other perceptual and motor modalities. Finally, the proposed method might find applications in hand-held wearable devices, or mobile phones, to improve patients' performance in the ADLs.

Chapter 9

AUTOMATED MULTISENSORY CUEING FACILITATES RECOVERY IN APHASIA: A TRIPLE-BLIND CLINICAL STUDY

This chapter is based on:
Grechuta, K., Ballester, B. R., Munné, R. E., Bernal, T. U., Hervás, B. M., Mohr, B., Pulvermuller, F., San Segundo, R., & Verschure, P. (2019) Automated multisensory cueing facilitates recovery in aphasia: a triple-blind clinical study. *Journal of NeuroEngineering and Rehabilitation.* (Under review)

Stroke can provoke serious language disorders which adversely impact quality of life and recovery of patients. Naming disorders, in particular, are the most frequent symptom in all types of aphasia. Thus the need for word-finding methods is pervasive both within and outside of the clinical context. Grounded in recent neurophysiological findings, in the reported study, we proposed and validated two novel multisensory cueing methods to aid word-finding difficulties using a virtual-reality based rehabilitation paradigm. The cues were integrated within a longitudinal clinical trial and presented in the form of semantically related sounds (“telephone” for the phone) or silent videos showing lips articulating the target word. Our findings provide clinical evidence for beneficial effects of the proposed multisensory cueing on naming and suggest that integrating such cues in the rehabilitation of aphasia could foster language-production skills even at the chronic stages of the disease.

9.1 Introduction

About 25% of stroke patients worldwide suffer from language disorders, and the majority remains chronic. Anomia, or word-finding difficulty, is a ubiquitous characteristic of aphasia which significantly compromises communication and quality of life of patients [Goodglass, 1980]. Consequently, aphasia rehabilitation largely incorporates strategies fostering the recovery of naming and verbal communication. Since word-finding is a multistep process, which consists of mapping a concept to its verbal structure [Foygel and Dell, 2000], naming therapies address different stages of retrieval. Phonologically oriented cueing targets the ability to retrieve phonemes underlying the articulation of a word, while semantic cueing targets the activation of lexical-semantic association networks [Levelt et al., 1999, Raymer et al., 2002]. Vast evidence demonstrates that both approaches improve immediate and long-term naming accuracy [Nickels, 2002] even if administered through technology-based methods [Abad et al., 2013, Kurland et al., 2014]. Grounded in the recent accounts of language processing [Pulvermüller and Fadiga, 2010, Kiefer and Pulvermüller, 2012], in the present study, we proposed two novel cueing techniques, Silent Visuomotor Cueing (SVC) and Semantic Auditory Cueing (SAC), and explored their effects on naming difficulties in 10 patients with chronic nonfluent aphasia. Cues were delivered within a longitudinal clinical intervention in which the participants underwent a dyadic Virtual Reality (VR)-based language therapy inspired by Intensive Language Action Therapy (ILAT) [Difrancesco et al., 2012]: the Rehabilitation Gaming System for aphasia (RGSa) [Grechuta et al., 2019a]. SACs contained auditory information semantically related to the target word (e.g., ringing sound for "telephone") while SVCs contained articulatory information presented in the form of a silent video [Grechuta et al., 2017a, Grechuta et al., 2019a]. We predicted that the proposed cues would facilitate word-finding and would result in faster dyadic interaction times.

9.2 Methods

9.2.1 Participants

Ten chronic patients ($Mean(SD)$: 69.9 ± 48.6 months post-stroke) with moderate-to-severe nonfluent aphasia participated in the study (age : $55 Mean(SD)$: 57.6 ± 9.8). All patients suffered a single left-hemispheric stroke. The reported paradigm followed a within-subjects design. The study was approved by the local ethics committee and registered on ClinicalTrials.gov (identifier: NCT02928822) [Grechuta et al., 2019a].

9.2.2 Protocol

The proposed cueing method was integrated within a novel language rehabilitation protocol, the Rehabilitation Gaming System for aphasia (RGSa) [Grechuta et al., 2019a]. RGSa is a VR-based dyadic rehabilitation tool inspired by ILAT [Difrancesco et al., 2012]. In the RGSa training, two patients engaged in an interactive language-training in the form of a turn-based game. Patients received five weekly intervention-sessions for two months. The goal of each session was to request and collect objects from the other player or hand them over when required. Patients interacted with the virtual objects by performing planar movements which were tracked and mapped onto avatar's arms. Both patients had a set of three identical objects simultaneously available for selection. Each trial consisted of three steps: (1)PlayerA chooses the desired object by reaching towards it, (2)PlayerA verbally requests it, and (3)PlayerB reacts to the request by reaching for the requested object and handing it over to Player A. After such successful trial, both patients received a point, heard the correct pronunciation of the target word and the turn changed (Figure 9.1 A). In every session, participants had to request and collect 36 objects.

9.2.3 Materials

The materials consisted of 120 3D objects. All items were categorized regarding frequency, semantic category, complexity, and phonemic similarity, and matched for the syllable length and semantic field. The materials were classified into two categories: items without semantically related sounds which in 50% of the trials underwent “Silent Visuomotor Cueing” (SVC, N=60, e.g. “bed”, “peach”, Figure 9.1Left), and those with semantically related sounds which in 50% of the trials underwent “Semantic Auditory Cueing” (SAC, N=60, e.g. “telephone”, “car”, Figure 9.1Right). SVCs consisted in displaying a silent video representing the movement of the lips correctly articulating the target word. SACs consisted of playing a sound semantically related to the object selected for the request. Both cues were provided immediately after object selection, once per trial. All participants were provided with a wireless headset through which they heard feedback from the system.

9.2.4 Measures

To evaluate learning of the target stimuli we administered the Vocabulary Test (VocabT) which included all the trained items [Grechuta et al., 2019a]. For each word, patients could score a maximum of 5 points (0:no verbal utterance, 1:utterance followed by full phonetic priming, 2:utterance followed by priming of the initial phoneme, 3:utterance followed by full silent orofacial hint, 4:utterance followed by silent orofacial hint of the first phoneme, 5:utterance followed by no hint). The test was administered six times over the intervention period (baseline, week 2, 4, 6, 8, 16). Finally, we extracted and computed ITs (Figure 9.1 A) for all stimuli and therapy sessions. We defined IT as the time interval between the selection and

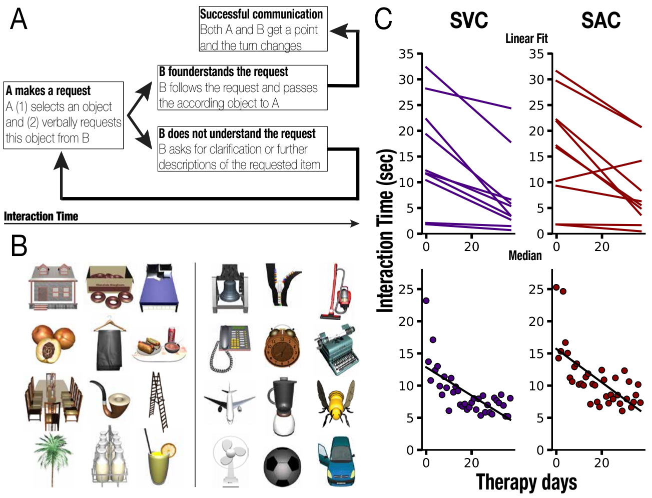


Figure 9.1: **(A)** Illustration of the Interaction Time (IT) measure, possible moves, and speech-acts. **(B)** Example of the materials. Left: stimuli undergoing SAC, right: stimuli undergoing SVC. **(C)** Fit for each patient's averaged IT over the therapy interval for all the stimuli undergoing Silent Visuomotor (SVC, violet) and Semantic Auditory (SAC, red) cueing. *Upper panels*: Lines represent linear regression models for individual patients including cued and non-cued trials. *Lower panels*: Median ITs of all the patients including all stimuli for each therapy session.

collection of an object. Each IT included word retrieval, articulation of the request, comprehension of the target word and the motor response of the opponent.

9.2.5 Data Analysis

We used the Wilcoxon signed-rank test to evaluate within-groups changes and Mann-Whitney U-test for between-group comparisons. All comparative analyses used two-tailed tests and a standard level of significance ($p < .05$).

9.3 Results

We aimed to determine the effects of SVC and SAC on lexical access and verbal execution in patients with chronic nonfluent aphasia. First, we evaluated whether ITs (Figure 9.1 A) reflect improvement in lexical production. Specifically, we examined the relationship between the proposed measure (IT) and verbal production evaluated on on a standard clinical scale (VocabT) which showed a significant increase from baseline after the intervention (Wilcoxon signed-rank: $p = .007$) [Grechuta et al., 2019a]. For the analysis, we extracted ITs including all cued and non-cued stimuli from all therapy sessions for each patient and computed the mean ITs collected on the date of the administration of the VocabT ± 1 day. Spearman’s correlation revealed a significant relationship between the two measures ($r = -.89, p = -.03$) suggesting that ITs may be regarded as a relevant measure of verbal execution.

The analysis of the evolution of the ITs throughout the intervention yielded a significant decrease for all the presented stimuli ($r = -.61, p < 0.001$) as well as for the two subsets chosen to undergo SVC (Figure 9.1 B Left, $r = -.7, p < 0.001$) and SAC (Figure 9.1 B Right, $r = -.69, p < 0.001$) (Figure 9.1 C) respectively. Subsequently, to estimate the effects of multisensory cues on verbal expression, we compared the ITs between cued and noncued stimuli for all the sessions as well as for the early and late trials (Figure 9.2 A). A Wilcoxon signed-rank test demonstrated a significant difference between all cued and noncued stimuli in SVC ($p = .001$) and SAC ($p = .003$). Specifically, we found that the difference between cued and noncued trials was statistically significant during the early therapy sessions (N=15) both for SVC ($p = .002$) and SAC ($p = .001$) (Figure 9.2 B). No differences in ITs were found in the late sessions for neither SVC ($p = .73$) nor SAC ($p = .53$) (Figure 9.2 B). Finally, the analysis yielded no differences in ITs between noncued SAC and SVC stimuli in the early sessions ($p = .28$) establishing that the chosen subsets did not differ regarding difficulty.

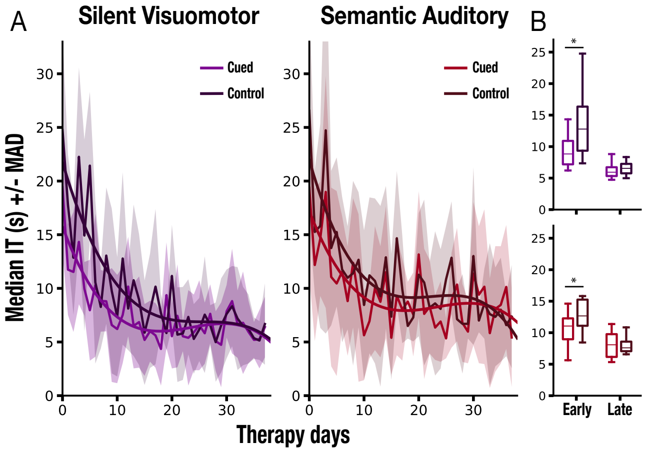


Figure 9.2: **The effects of multisensory cueing on the Interaction Times (IT). (A)** Evolution of median ITs for cued on noncued stimuli over the therapy sessions. Lines represent nonlinear regression models for cued and noncued visuomotor (violet) and auditory (red) cues. **(B)** Quantification of differences in ITs for SVC and SAC between cued and noncued stimuli in the early (first 15) and late (last 15) therapy sessions.

9.4 Discussion

This study aimed to examine the effects of multisensory cueing on verbal execution in patients with chronic nonfluent aphasia using a VR-based dyadic language-rehabilitation protocol, the RGSa [Grechuta et al., 2016a, Grechuta et al., 2019a]. The system allowed for automatic, continuous quantification of the priming effects within a longitudinal clinical trial [Grechuta et al., 2019a]. Specifically, for all intervention sessions (38 days) and each patient ($N = 10$), we stored the Interaction Times (IT, Figure 9.1 A), which included the explicit selection of a target item, word-finding, and the opponent’s motor response. ITs significantly decreased over the time of the intervention (Figure 9.1 C). Crucially, the proposed measure was strongly correlated with the scores on a standard clinical scale evaluating naming (VocabT) suggesting that ITs might be regarded as a reliable diagnostic tool for changes in word production, which does not require assistance or supervision of a therapist.

ITs were significantly lower in cued vs. noncued trials for both Semantic Auditory (SAC) and Silent Visuomotor (SVC) stimuli sets (Figure 9.2 A). The subsequent analyses revealed differences between cued and noncued trials for SAC and SVC in the early intervention days when the exposure to the target lexicon was still infrequent (Figure 9.2 B). No such differences occurred during late sessions when the acquisition of the target stimuli might have reached a plateau (Figure 9.2 B). These results provide clinical evidence for beneficial effects of multisensory cueing on verbal execution and suggest that integrating SVC and SAC in the rehabilitation of aphasia could foster language-production skills within and outside of the clinic even at the chronic stages of the disease.

Our findings are consistent with sensorimotor accounts of language processing [Pulvermüller and Fadiga, 2010, Kiefer and Pulvermüller, 2012, Massaro and Palmer, 1998] highlighting the tight neuronal connection between language and motor brain regions. In particular, SVCs might have triggered the activation of the language networks related to the processing of orofacial gestures facilitating articulation [Petrides et al., 2005, Nishitani and Hari, 2002], while SACs could have prompted the activation of semantic regions facilitating lexical access [Kiefer et al., 2008]. Future studies shall systematically investigate the generalization effects of the proposed cueing technique beyond the clinical context. At the current stage, the proposed method might be integrated into wearable or mobile technologies as well as virtual reality based systems to improve patients' language performance in daily life.

Chapter 10

CONCLUSIONS

This dissertation aimed to shed light on the question of Alice which asks "Who in the world am I?" and advance our understanding of the principles underlying lower-level processes and higher-level functions in the context of goal-oriented action. In particular, we grounded our approach in the theoretical and empirical framework of embodied cognition, also referred to as the sensorimotor approach, which critically challenges the traditional view at the basis of cognition including body ownership and language learning [Merleau-Ponty, 1945, Barsalou, 2008, Verschure, 2011, Pulvermüller and Fadiga, 2010, Kiefer and Pulvermüller, 2012, Massaro and Palmer, 1998]. In doing so, we employed a series of experiments using bodily illusions, on the one hand, and clinical trials with patients who suffer nonfluent aphasia, on the other. Altogether, our results shed light on how lower-level processes, such as the feeling of the body as belonging to oneself, and higher-level cognitive functions, such as language (re)learning are tightly coupled to the motor system, goals at hand and the dynamics of the surrounding environment. Crucially, this work might broaden the understating of body ownership as well as approaches to language rehabilitation, which would not be possible if a traditional disembodied method, which does not consider the sensorimotor foundations of learning, was applied.

In the following sections, we will draw the main conclusions in the context of the present work dedicated to the experience of body ownership and the rehabilitation of language, respectively. Finally, we will sketch an outlook for the future work comprising alternative methods which shall be used to support current results from the neural perspective, as well as introduce intriguing questions which still require empirical evidence.

10.1 Insights on Body Ownership

Different from classical theories, which view the body as functionally independent from any type of cognition, contemporary psychology, and cognitive neuroscience has drawn much attention to the notion of the bodily self and the study of its biological, cognitive, and philosophical basis [Gallagher, 2006, O’Regan and Noë, 2001, Noë, 2004]. Within this framework, the concept of *body ownership* which allows us to determine the boundaries of the physical self thus enabling subjective experience has gained much attention in the field of experimental psychology [Tsakiris, 2010, Blanke, 2012]. Grounded in the theoretical framework of embodiment as well as recent experiment-driven insights from cognitive science, which support the notion that self-awareness is inseparable from bodily activities and sensorimotor expectations [O’Regan and Noë, 2001, Noë, 2004, Verschure et al., 2003], in the present dissertation we aimed at addressing some of the critical questions which regard the principles of body ownership in the context of self-generated action. This work, comprised of four studies, was presented in Part I of this thesis.

First (Chapter 2), we asked the question of whether the embodiment hypothesis about the coupling between the body and environment through action as well as the structural overlap of cortical regions underlying the experience of ownership and motor planning is of any functional significance [Grechuta et al., 2017b]. The goal of the second study (Chapter 3) was to examine whether body ownership, coupled to the motor system, depends on not only the bodily cues but also purely distal sensory signals which pertain to a goal-oriented task, similar to a forward model [Grechuta et al., 2019d]. Inspired by the results from the second contribution as well as other studies [Ma and Hommel, 2015b], in the third experiment (Chapter 4) we prompted the notions of connectedness, peripersonal space, and physical plausibility of the effector as necessary conditions for the experience of ownership in the context of action [De la Torre et al., 2019]. And finally, our goal was to explore the effects of the congruence of the external world on the feeling of ownership (Chapter 5) [Grechuta et al., 2019c]. In line with the embodiment thesis which emphasizes functional coupling between the brain, body, and environment, we proposed that body ownership might also depend on the integration and prediction of purely distal sensory signals which pertain to the environment and violate our prior knowledge.

The results from the set of those four experiments shed new light on the principles of body ownership. Specifically, from the first study, we learned that body ownership indeed seems to be coupled to the systems of motor control and decision making. Moreover, the obtained results might suggest that visuotactile associations possibly generalize into a coherent percept of a body through action. Indeed, even in the classical Rubber Hand Illusion [Botvinick and Cohen, 1998], one can observe behavior similar to a withdrawal reflex when the rubber hand is threatened by the knife. This reaction can act as a proxy to suggest that, although the participants cannot move the rubber hand, the ownership which they experience over it might be constructing the underlying experience of the agency. The second study further demonstrated the coupling between the motor system and the experience of ownership of a virtual body. Specifically, we showed that body ownership is affected by incongruencies of purely distal modalities (i.e., visual and auditory) which pertain to a goal-oriented task and which are possibly computed by a forward model or corollary discharge, a similar mechanism to the one modulating agency. These results challenge the traditional boundaries of the physical self, and are in line with the hypothesis of Merleau-Ponty [Merleau-Ponty, 1945] who proposed that in the context of action our body seems to act like a *vehicle*, thus depending on all the sensory cues which inform about the outcomes of the task at hand independently of their source. This finding led us to the perform the third study in which we showed that the *vehicle* can actually be any effector which does not require to be connected with the real body and which can be located outside of the traditionally understood peripersonal space [Maravita and Iriki, 2004] provided that it can be fully controlled and that its behavior is entirely predictable. Interestingly, similar to the previous studies, here we found support for the close link between the notion of body ownership and agency.

The goal of the final contribution within the scope of Part I of this thesis was to understand the role of the environment in the brain-body-environment coupling at the basis of ownership. In particular, we aimed at investigating what is the contribution of the prior knowledge about the rules of the environment to the experience of one's body while performing a goal-oriented action? Curiously, it seems that the incongruencies in action-independent and task-irrelevant sensory cues, which inform about the statistical structure of the environment and are processed by purely distal modalities, modulate the experience of body ownership. This suggests that similarly to any robust precept, body ownership depends on the consistency of the internal models of not only the body or the consequences of its actions but also the model of the surrounding environment. Crucially the modulation of ownership was not as strong as either in the classical Rubber Hand Illusion, or when the action-driven cues are manipulated. Together with what we have known about body ownership from classical studies, our findings might suggest that there is a gradient of sensory predictions stemming from the introspective signals [Seth, 2013] to bodily [Tsakiris, 2010, Blanke, 2012] and exteroceptive cues [Grechuta et al., 2019d, De la Torre et al., 2019] which weight ownership in a hierarchical fashion, depending on the current goal- a hypothesis which shall be investigated in future studies.

Overall, the studies reported in the first part of this thesis elucidate the first dimension of the opening question and broaden our understating of self-awareness by complementing previous models of body ownership grounded in the Rubber Hand Illusion paradigm [Tsakiris, 2010, Blanke, 2012, Botvinick, 2004, Guterstam et al., 2019]. It is worth to note, however, that while the previous studies investigated the contribution of purely exafferent (i.e., tactile) cues, in the present work we focused on action-driven reafferent signals. These two approaches seem fundamentally different. Specifically, the classical view assumes that there exists such a sensory state in which an organism can only passively receive sensory stimuli, almost like having a brain placed in a vat and connected to external input sources [Dennett, 2017]. Conversely, the present approach was designed under an assumption that the function of the brain is to control the behavior of a goal-oriented individual whose objective in to survive. Future studies shall bridge the methodological gap between these two conceptually different approaches to fully understand what stands at the basis of the bodily self in the context when both reafferent and exafferent cues are present.

10.2 Insights on Language (re)Learning

Inspired by the theories of embodied cognition, we have found converging empirical evidence which supports the notion that sensory and motor systems play a critical role in higher-level cognitive functions such as the construal of meaning and language processing [Barsalou, 2008, Lakoff and Johnson, 2008, Pulvermüller, 2005, Glenberg and Kaschak, 2002]. On the one hand, these findings are relevant for the advancement of basic neuroscience and the understanding of the basis of linguistic processing. On the other hand, however, as we discussed in the motivation of this thesis, they also critically challenge traditional approaches to, among others, language rehabilitation. The second goal of this dissertation was to shed light on the second dimension of Alice's question by explicitly studying whether the principles of embodiment grounded in neuroscientific evidence and integrated within a multimodal dyadic language therapy could positively impact (re)learning of the linguistic skill and long-term retention of the changes. Moreover, grounded in the empirical evidence of embodied cognition about the multimodal grounding of concepts, we aimed at understating whether providing multisensory cues (e.g., auditory or visual) at the moment of word retrieval could aid lexical access and verbal production during the RGSs sessions.

First, we designed and implemented a virtual-reality based language rehabilitation in the form of a linguistic game inspired by Intensive Language-Action Therapy (ILAT) [Difrancesco et al., 2012]. In the initial pilot study (Chapter 6), we aimed at evaluating the usability and safety of the proposed RGSa system. Based on the promising results of the first study, we designed a longitudinal Randomized Control Trial (RTC) to systematically evaluate the effectiveness of the RGSa as compared to the standard speech and language therapy in seventeen post-stroke patients with persistent nonfluent aphasia (Chapter 7) [Grechuta et al., 2019a]. Within the same clinical trial we tested the hypothesis which is reported in the third (Chapter 8) [Grechuta et al., 2017a] and fourth (Chapter 9) [Grechuta et al., 2019b] study. In particular, we hypothesized that if, as proposed by the embodiment hypothesis, mental representations of concepts are multimodal constituting multiple sensory and motor features, then providing any sensory cue associated to the target word would facilitate word retrieval in patients with chronic nonfluent aphasia.

Altogether, the results from the clinical work presented in Part II of this thesis had two main outcomes. First, in line with other recent studies [Pulvermüller and Berthier, 2008, Kurland et al., 2012, Kurland et al., 2018, Stahl et al., 2016, Berthier and Pulvermüller, 2011], we have demonstrated that a behaviourally relevant and goal-oriented treatment based on the principles that language and concepts are *grounded* [Zwaan et al., 2002, Stanfield and Zwaan, 2001, Connell, 2007, Glenberg and Robertson, 2000, Gerlach et al., 2002, Kaschak and Glenberg, 2000, Spence et al., 2001, Borghi et al., 2004, Willems et al., 2010, Lyons et al., 2010, Pulvermüller, 2005] positively impacts recovery of aphasia and long-term stability of the effects. Secondly, our work supports the notion that the challenges faced by the health-care system in the treatment of post-stroke symptoms might be effectively addressed by augmenting traditional therapy with principle-based computer methods.

Bibliography

[Abad et al., 2013] Abad, A., Pompili, A., Costa, A., Trancoso, I., Fonseca, J., Leal, G., Farrajota, L., and Martins, I. P. (2013). Automatic word naming recognition for an on-line aphasia treatment system. *Computer Speech & Language*, 27(6):1235–1248.

[Abel et al., 2007] Abel, S., Willmes, K., and Huber, W. (2007). Model-oriented naming therapy: Testing predictions of a connectionist model. *Aphasiology*, 21(5):411–447.

[Angela and Dayan, 2005] Angela, J. Y. and Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron*, 46(4):681–692.

[Antle et al., 2009] Antle, A. N., Corness, G., Bakker, S., Droumeva, M., van den Hoven, E., and Bevans, A. (2009). Designing to support reasoned imagination through embodied metaphor. In *Proceedings of the seventh ACM conference on Creativity and cognition*, pages 275–284. ACM.

[Apps and Tsakiris, 2014] Apps, M. A. and Tsakiris, M. (2014). The free-energy self: a predictive coding account of self-recognition. *Neuroscience & Biobehavioral Reviews*, 41:85–97.

[Arbib, 2010] Arbib, M. A. (2010). Mirror system activity for action and language is embedded in the integration of dorsal and ventral pathways. *Brain and language*, 112(1):12–24.

[Arévalo et al., 2007] Arévalo, A., Perani, D., Cappa, S., Butler, A., Bates, E., and Dronkers, N. (2007). Action and object processing in aphasia: From nouns and verbs to the effect of manipulability. *Brain and language*, 100(1):79–94.

[Armel and Ramachandran, 2003] Armel, K. C. and Ramachandran, V. S. (2003). Projecting sensations to external objects: evidence from skin conductance response. *Proceedings of the Royal Society of London B: Biological Sciences*, 270(1523):1499–1506.

[Baier and Karnath, 2008] Baier, B. and Karnath, H.-O. (2008). Tight link between our sense of limb ownership and self-awareness of actions. *Stroke*, 39(2):486–488.

[Bak and Hodges, 2004] Bak, T. H. and Hodges, J. R. (2004). The effects of motor neurone disease on language: further evidence. *Brain and language*, 89(2):354–361.

[Bak et al., 2005] Bak, T. H., Yancopoulou, D., Nestor, P. J., Xuereb, J. H., Spillantini, M. G., Pulvermüller, F., and Hodges, J. R. (2005). Clinical, imaging and pathological correlates of a hereditary deficit in verb and action processing. *Brain*, 129(2):321–332.

[Ballester et al., 2015a] Ballester, B. R., Maier, M., San Segundo, R., Galeano, V. C., Duff, A., and Verschure, P. F. (2015a). Reinforcement-induced movement therapy: A novel approach for overcoming learned non-use in chronic stroke patients. In *Virtual Rehabilitation Proceedings (ICVR), 2015 International Conference on*, pages 183–190. IEEE.

[Ballester et al., 2015b] Ballester, B. R., Nirme, J., Duarte, E., Cuxart, A., Rodriguez, S., Verschure, P., and Duff, A. (2015b). The visual amplification of goal-oriented movements counteracts acquired non-use in hemiparetic stroke patients. *Journal of neuroengineering and rehabilitation*, 12(1):50.

[Ballester et al., 2015c] Ballester, B. R., Oliva, L. S., Duff, A., and Verschure, P. (2015c). Accelerating motor adaptation by virtual reality based modulation of error memories. In *Rehabilitation Robotics (ICORR), 2015 IEEE International Conference on*, pages 623–629. IEEE.

[Banakou et al., 2013] Banakou, D., Groten, R., and Slater, M. (2013). Illusory ownership of a virtual child body causes overestimation of object sizes and implicit attitude changes. *Proceedings of the National Academy of Sciences*, 110(31):12846–12851.

[Banks et al., 1989] Banks, G., Short, P., Martínez, A. J., Latchaw, R., Ratcliff, G., and Boller, F. (1989). The alien hand syndrome: clinical and postmortem findings. *Archives of Neurology*, 46(4):456–459.

[Banzi et al., 2013] Banzi, M., Cuartielles, D., Igoe, T., Martino, G., and Mellis, D. (2013). Site arduino.

[Barnsley et al., 2011] Barnsley, N., McAuley, J., Mohan, R., Dey, A., Thomas, P., and Moseley, G. (2011). The rubber hand illusion increases histamine reactivity in the real arm. *Current Biology*, 21(23):R945–R946.

[Barsalou, 1999] Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and brain sciences*, 22(4):577–660.

[Barsalou, 2008] Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.*, 59:617–645.

[Barsalou et al., 2003] Barsalou, L. W., Simmons, W. K., Barbey, A. K., and Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in cognitive sciences*, 7(2):84–91.

[Barsalou and Wiemer-Hastings, 2005] Barsalou, L. W. and Wiemer-Hastings, K. (2005). Situating abstract concepts. *Grounding cognition: The role of perception and action in memory, language, and thought*, pages 129–163.

[Bäß et al., 2008] Bäß, P., Jacobsen, T., and Schröger, E. (2008). Suppression of the auditory n1 event-related potential component with unpredictable self-initiated tones: evidence for internal forward models with dynamic stimulation. *International Journal of Psychophysiology*, 70(2):137–143.

[Beck et al., 2008] Beck, J. M., Ma, W. J., Kiani, R., Hanks, T., Churchland, A. K., Roitman, J., Shadlen, M. N., Latham, P. E., and Pouget, A. (2008). Probabilistic population codes for bayesian decision making. *Neuron*, 60(6):1142–1152.

[Beilock et al., 2008] Beilock, S. L., Lyons, I. M., Mattarella-Micke, A., Nusbaum, H. C., and Small, S. L. (2008). Sports experience changes the neural processing of action language. *Proceedings of the National Academy of Sciences*, 105(36):13269–13273.

[Benedek and Kaernbach, 2010a] Benedek, M. and Kaernbach, C. (2010a). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, 190(1):80–91.

[Benedek and Kaernbach, 2010b] Benedek, M. and Kaernbach, C. (2010b). Decomposition of skin conductance data by means of nonnegative deconvolution. *Psychophysiology*, 47(4):647–658.

[Berthier and Pulvermüller, 2011] Berthier, M. L. and Pulvermüller, F. (2011). Neuroscience insights improve neurorehabilitation of poststroke aphasia. *Nature Reviews Neurology*, 7(2):86.

[Berti et al., 2005] Berti, A., Bottini, G., Gandola, M., Pia, L., Smania, N., Stracciari, A., Castiglioni, I., Vallar, G., and Paulesu, E. (2005). Shared cortical anatomy for motor awareness and motor control. *Science*, 309(5733):488–491.

[Blakemore et al., 2000] Blakemore, S.-J., Wolpert, D., and Frith, C. (2000). Why can't you tickle yourself? *Neuroreport*, 11(11):R11–R16.

[Blanke, 2012] Blanke, O. (2012). Multisensory brain mechanisms of bodily self-consciousness. *Nature Reviews Neuroscience*, 13(8):556–571.

[Bloom and Kirsch, 2016] Bloom, A. and Kirsch, A. (2016). *The republic of Plato*. Basic Books.

[Boo and Rose, 2011] Boo, M. and Rose, M. L. (2011). The efficacy of repetition, semantic, and gesture treatments for verb retrieval and use in broca's aphasia. *Aphasiology*, 25(2):154–175.

[Borghi et al., 2004] Borghi, A. M., Glenberg, A. M., and Kaschak, M. P. (2004). Putting words in perspective. *Memory & Cognition*, 32(6):863–873.

[Borreggine and Kaschak, 2006] Borreggine, K. L. and Kaschak, M. P. (2006). The action–sentence compatibility effect: It's all in the timing. *Cognitive Science*, 30(6):1097–1112.

[Botvinick, 2004] Botvinick, M. (2004). Probing the neural basis of body ownership. *Science*, 305(5685):782–783.

[Botvinick and Cohen, 1998] Botvinick, M. and Cohen, J. (1998). Rubber hands' feel'touch that eyes see. *Nature*, 391(6669):756.

[Boulenger et al., 2008a] Boulenger, V., Mechtouff, L., Thobois, S., Broussolle, E., Jeannerod, M., and Nazir, T. A. (2008a). Word processing in parkinson's disease is impaired for action verbs but not for concrete nouns. *Neuropsychologia*, 46(2):743–756.

[Boulenger et al., 2006] Boulenger, V., Roy, A. C., Paulignan, Y., Deprez, V., Jeannerod, M., and Nazir, T. A. (2006). Cross-talk between language processes and overt motor behavior in the first 200 msec of processing. *Journal of cognitive neuroscience*, 18(10):1607–1615.

[Boulenger et al., 2008b] Boulenger, V., Silber, B. Y., Roy, A. C., Paulignan, Y., Jeannerod, M., and Nazir, T. A. (2008b). Subliminal display of action words interferes with motor planning: a combined eeg and kinematic study. *Journal of Physiology-Paris*, 102(1-3):130–136.

[Bourbonnais and Noven, 1989] Bourbonnais, D. and Noven, S. V. (1989). Weakness in patients with hemiparesis. *American Journal of Occupational Therapy*, 43(5):313–319.

[Brown and Brüne, 2012] Brown, E. C. and Brüne, M. (2012). The role of prediction in social neuroscience. *Frontiers in human neuroscience*, 6:147.

[Brozzoli et al., 2011a] Brozzoli, C., Gentile, G., Petkova, V. I., and Ehrsson, H. H. (2011a). Fmri adaptation reveals a cortical mechanism for the coding of space near the hand. *Journal of Neuroscience*, 31(24):9023–9031.

[Brozzoli et al., 2011b] Brozzoli, C., Makin, T. R., Cardinali, L., Holmes, N. P., and Farne, A. (2011b). *Peripersonal space: a multisensory interface for body-object interactions*. Taylor & Francis.

[Bruce and Howard, 1987] Bruce, C. and Howard, D. (1987). Computer-generated phonemic cues: an effective aid for naming in aphasia. *International Journal of Language & Communication Disorders*, 22(3):191–201.

[Buccino et al., 2005] Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V., and Rizzolatti, G. (2005). Listening to action-related sentences modulates the activity of the motor system: a combined tms and behavioral study. *Cognitive Brain Research*, 24(3):355–363.

[Buxbaum and Saffran, 2002] Buxbaum, L. J. and Saffran, E. M. (2002). Knowledge of object manipulation and object function: dissociations in apraxic and nonapraxic subjects. *Brain and language*, 82(2):179–199.

[Cameirao et al., 2012] Cameirao, M. S., i Badia, S. B., Duarte, E., Frisoli, A., and Verschure, P. F. (2012). The combined impact of virtual reality neurorehabilitation and its interfaces on upper extremity functional recovery in patients with chronic stroke. *Stroke*, 43(10):2720–2728.

[Cameirão et al., 2010] Cameirão, M. S., i Badia, S. B., Oller, E. D., and Verschure, P. F. (2010). Neurorehabilitation using the virtual reality based rehabilitation gaming system: methodology, design, psychometrics, usability and validation. *Journal of neuroengineering and rehabilitation*, 7(1):48.

[Cameirão et al., 2007] Cameirão, M. S., i Badia, S. B., Zimmerli, L., Oller, E. D., and Verschure, P. F. (2007). The rehabilitation gaming system: a virtual reality based system for the evaluation and rehabilitation of motor deficits. In *Virtual Rehabilitation, 2007*, pages 29–33. IEEE.

[Cao et al., 1999] Cao, Y., Vikingstad, E. M., George, K. P., Johnson, A. F., and Welch, K. (1999). Cortical language activation in stroke patients recovering from aphasia with functional mri. *Stroke*, 30(11):2331–2340.

[Carlsson et al., 2000] Carlsson, K., Petrovic, P., Skare, S., Petersson, K. M., and Ingvar, M. (2000). Tickling expectations: neural processing in anticipation of a sensory stimulus. *Journal of cognitive neuroscience*, 12(4):691–703.

[Carter et al., 2012] Carter, A. R., Shulman, G. L., and Corbetta, M. (2012). Why use a connectivity-based approach to study stroke and recovery of function? *Neuroimage*, 62(4):2271–2280.

[Castiello et al., 1991] Castiello, U., Paulignan, Y., and Jeannerod, M. (1991). Temporal dissociation of motor responses and subjective awareness. *Brain*, 114(6):2639–2655.

[Chalmers, 1993] Chalmers, D. J. (1993). A computational foundation for the study of cognition.

[Clark, 1999] Clark, A. (1999). An embodied cognitive science? *Trends in cognitive sciences*, 3(9):345–351.

[Clark, 2013] Clark, A. (2013). Whatever next? predictive brains, situated agents, and the future of cognitive science. *Behavioral and brain sciences*, 36(3):181–204.

[Cole and Cole, 1995] Cole, J. and Cole, J. O. (1995). *Pride and a daily marathon*. MIT Press.

[Collins and Loftus, 1975] Collins, A. M. and Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological review*, 82(6):407.

[Connell, 2007] Connell, L. (2007). Representing object colour in language comprehension. *Cognition*, 102(3):476–485.

[Corballis, 2009] Corballis, M. C. (2009). Language as gesture. *Human Movement Science*, 28(5):556–565.

[Coslett, 1998] Coslett, H. B. (1998). Evidence for a disturbance of the body schema in neglect. *Brain and cognition*, 37(3):527–544.

[Costantini, 2014] Costantini, M. (2014). Body perception, awareness, and illusions. *Wiley Interdisciplinary Reviews: Cognitive Science*, 5(5):551–560.

[Costantini and Haggard, 2007] Costantini, M. and Haggard, P. (2007). The rubber hand illusion: sensitivity and reference frame for body ownership. *Consciousness and cognition*, 16(2):229–240.

[Craft et al., 2005] Craft, T. K., Glasper, E. R., McCullough, L., Zhang, N., Sugo, N., Otsuka, T., Hurn, P. D., and DeVries, A. C. (2005). Social interaction improves experimental stroke outcome. *Stroke*, 36(9):2006–2011.

[Craig, 2002] Craig, A. D. (2002). How do you feel? interoception: the sense of the physiological condition of the body. *Nature reviews neuroscience*, 3(8):655–666.

[Craig and Craig, 2009] Craig, A. D. and Craig, A. (2009). How do you feel–now? the anterior insula and human awareness. *Nature reviews neuroscience*, 10(1).

[Crapse and Sommer, 2008] Crapse, T. B. and Sommer, M. A. (2008). Corollary discharge across the animal kingdom. *Nature Reviews Neuroscience*, 9(8):587.

[Damasio and Geschwind, 1984] Damasio, A. R. and Geschwind, N. (1984). The neural basis of language. *Annual review of neuroscience*, 7(1):127–147.

[Damasio and Macaluso, 2004] Damasio, A. R. and Macaluso, F. (2004). *L'errore di Cartesio: emozione, ragione e cervello umano*. Adelphi.

[D'Angelo et al., 2018] D'Angelo, M., di Pellegrino, G., Seriani, S., Gallina, P., and Frassinetti, F. (2018). The sense of agency shapes body schema and peripersonal space. *Scientific reports*, 8.

[D'Ausilio et al., 2009] D'Ausilio, A., Pulvermüller, F., Salmas, P., Bufalari, I., Begliomini, C., and Fadiga, L. (2009). The motor somatotopy of speech perception. *Current Biology*, 19(5):381–385.

[David et al., 2016] David, N., Skoruppa, S., Gulberti, A., Schultz, J., and Engel, A. K. (2016). The sense of agency is more sensitive to manipulations of outcome than movement-related feedback irrespective of sensory modality. *PloS one*, 11(8):e0161156.

[Davis and Johnsrude, 2003] Davis, M. H. and Johnsrude, I. S. (2003). Hierarchical processing in spoken language comprehension. *Journal of Neuroscience*, 23(8):3423–3431.

[De Beauvoir, 1945] De Beauvoir, S. (1945). *La phénoménologie de la perception de Maurice Merleau-Ponty*.

[De la Torre et al., 2019] De la Torre, J., Grechuta, K., Daversa, D., and Verschure, P. (2019). (Body) ownership as a matter of control, not priors, connectedness or peripersonal space. *In preparation*.

[De Vignemont, 2014] De Vignemont, F. (2014). A multimodal conception of bodily awareness. *Mind*, 123(492):989–1020.

[Dennett, 2017] Dennett, D. C. (2017). *Brainstorms: Philosophical essays on mind and psychology*. MIT press.

[Descartes, 1958] Descartes, R. (([1637]1958)). 1637/1958: Dioptric. *New York: Modern Library*.

[Devue et al., 2007] Devue, C., Collette, F., Balteau, E., Degueldre, C., Luxen, A., Maquet, P., and Brédart, S. (2007). Here i am: the cortical correlates of visual self-recognition. *Brain research*, 1143:169–182.

[Dieguez et al., 2009] Dieguez, S., Mercier, M. R., Newby, N., and Blanke, O. (2009). Feeling numbness for someone else's finger. *Current Biology*, 19(24):R1108–R1109.

[Difrancesco et al., 2012] Difrancesco, S., Pulvermüller, F., and Mohr, B. (2012). Intensive language-action therapy (ilat): The methods. *Aphasiology*, 26(11):1317–1351.

[Ding et al., 2003] Ding, Y., Li, J., Clark, J., Diaz, F. G., and Rafols, J. (2003). Synaptic plasticity in thalamic nuclei enhanced by motor skill training in rat with transient middle cerebral artery occlusion. *Neurological Research*, 25(2):189–194.

[Doya et al., 2007] Doya, K., Ishii, S., Pouget, A., and Rao, R. P. (2007). *Bayesian brain: Probabilistic approaches to neural coding*. MIT press.

[Dummer et al., 2009] Dummer, T., Picot-Annand, A., Neal, T., and Moore, C. (2009). Movement and the rubber hand illusion. *Perception*, 38(2):271–280.

[Edmonds et al., 2009] Edmonds, L. A., Nadeau, S. E., and Kiran, S. (2009). Effect of verb network strengthening treatment (vnest) on lexical retrieval of content words in sentences in persons with aphasia. *Aphasiology*, 23(3):402–424.

[Ehrsson, 2012] Ehrsson, H. H. (2012). *The Concept of Body Ownership and Its Relation to Multisensory Integration*. MIT Press Cambridge, MA.

[Ehrsson et al., 2005] Ehrsson, H. H., Holmes, N. P., and Passingham, R. E. (2005). Touching a rubber hand: feeling of body ownership is associated with activity in multisensory brain areas. *The Journal of Neuroscience*, 25(45):10564–10573.

[Ehrsson et al., 2004] Ehrsson, H. H., Spence, C., and Passingham, R. E. (2004). That's my hand! activity in premotor cortex reflects feeling of ownership of a limb. *Science*, 305(5685):875–877.

[Ehrsson et al., 2007] Ehrsson, H. H., Wiech, K., Weiskopf, N., Dolan, R. J., and Passingham, R. E. (2007). Threatening a rubber hand that you feel is yours elicits a cortical anxiety response. *Proceedings of the National Academy of Sciences*, 104(23):9828–9833.

[Elman et al., 2000] Elman, R. J., Ogar, J., and Elman, S. H. (2000). Aphasia: Awareness, advocacy, and activism. *Aphasiology*, 14(5-6):455–459.

[Engel et al., 2001] Engel, A. K., Fries, P., and Singer, W. (2001). Dynamic predictions: oscillations and synchrony in top–down processing. *Nature Reviews Neuroscience*, 2(10):704.

[Ernst and Bülthoff, 2004] Ernst, M. O. and Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in cognitive sciences*, 8(4):162–169.

[Fadiga et al., 2002] Fadiga, L., Craighero, L., Buccino, G., and Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: a tms study. *European Journal of Neuroscience*, 15(2):399–402.

[Fadiga et al., 2009] Fadiga, L., Craighero, L., and D'Ausilio, A. (2009). Broca's area in language, action, and music. *Annals of the New York Academy of Sciences*, 1169(1):448–458.

[Fadiga et al., 1995] Fadiga, L., Fogassi, L., Pavesi, G., and Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *Journal of neurophysiology*, 73(6):2608–2611.

[Farmer et al., 2012] Farmer, H., Tajadura-Jiménez, A., and Tsakiris, M. (2012). Beyond the colour of my skin: How skin colour affects the sense of body-ownership. *Consciousness and cognition*, 21(3):1242–1256.

[Farnè and Làdavas, 2000] Farnè, A. and Làdavas, E. (2000). Dynamic size-change of hand peripersonal space following tool use. *Neuroreport*, 11(8):1645–1649.

[Farrer et al., 2003] Farrer, C., Franck, N., Georgieff, N., Frith, C. D., Decety, J., and Jeannerod, M. (2003). Modulating the experience of agency: a positron emission tomography study. *Neuroimage*, 18(2):324–333.

[Farrer and Frith, 2002] Farrer, C. and Frith, C. D. (2002). Experiencing oneself vs another person as being the cause of an action: the neural correlates of the experience of agency. *Neuroimage*, 15(3):596–603.

[Feinberg, 2002] Feinberg, T. E. (2002). *Altered egos: How the brain creates the self.* Oxford University Press.

[Ferri et al., 2017] Ferri, F., Ambrosini, E., Pinti, P., Merla, A., and Costantini, M. (2017). The role of expectation in multisensory body representation–neural evidence. *European Journal of Neuroscience*, 46(3):1897–1905.

[Ferri et al., 2013] Ferri, F., Chiarelli, A. M., Merla, A., Gallese, V., and Costantini, M. (2013). The body beyond the body: expectation of a sensory event is enough to induce ownership over a fake hand. *Proceedings of the Royal Society of London B: Biological Sciences*, 280(1765):20131140.

[Ferri and Costantini, 2016] Ferri, F. and Costantini, M. (2016). Commentary: The magnetic touch illusion: A perceptual correlate of visuo-tactile integration in peripersonal space. *Frontiers in human neuroscience*, 10:492.

[Fink et al., 2005] Fink, R., Brecher, A., Sobel, P., and Schwartz, M. (2005). Computer-assisted treatment of word retrieval deficits in aphasia. *Aphasiology*, 19(10-11):943–954.

[Fodor, 1975] Fodor, J. A. (1975). *The language of thought*, volume 5. Harvard university press.

[Fodor, 1983] Fodor, J. A. (1983). *The modularity of mind.* MIT press.

[Foltz et al., 1998] Foltz, P. W., Kintsch, W., and Landauer, T. K. (1998). The measurement of textual coherence with latent semantic analysis. *Discourse processes*, 25(2-3):285–307.

[Fotopoulou et al., 2011] Fotopoulou, A., Jenkinson, P. M., Tsakiris, M., Haggard, P., Rudd, A., and Kopelman, M. D. (2011). Mirror-view reverses somatoparaphrenia: dissociation between first-and third-person perspectives on body ownership. *Neuropsychologia*, 49(14):3946–3955.

[Foygel and Dell, 2000] Foygel, D. and Dell, G. S. (2000). Models of impaired lexical access in speech production. *Journal of Memory and Language*, 43:182–216.

[Fridriksson et al., 2012] Fridriksson, J., Hubbard, H. I., Hudspeth, S. G., Holland, A. L., Bonilha, L., Fromm, D., and Rorden, C. (2012). Speech entrainment enables patients with broca's aphasia to produce fluent speech. *Brain*, 135(12):3815–3829.

[Friston, 2010] Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11(2):127–138.

[Friston, 2012] Friston, K. (2012). Prediction, perception and agency. *International Journal of Psychophysiology*, 83(2):248–252.

[Friston et al., 2006] Friston, K., Kilner, J., and Harrison, L. (2006). A free energy principle for the brain. *Journal of Physiology-Paris*, 100(1-3):70–87.

[Friston et al., 2010] Friston, K. J., Daunizeau, J., Kilner, J., and Kiebel, S. J. (2010). Action and behavior: a free-energy formulation. *Biological cybernetics*, 102(3):227–260.

[Frith, 2013] Frith, C. (2013). *Making up the mind: How the brain creates our mental world.* John Wiley & Sons.

[Gallagher, 1986] Gallagher, S. (1986). Lived body and environment. *Research in phenomenology*, 16(1):139–170.

[Gallagher, 2000] Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive science. *Trends in cognitive sciences*, 4(1):14–21.

[Gallagher, 2006] Gallagher, S. (2006). *How the body shapes the mind.* Clarendon Press.

[Gallagher, 2007] Gallagher, S. (2007). The natural philosophy of agency. *Philosophy Compass*, 2(2):347–357.

[Gallagher, 2008] Gallagher, S. (2008). Direct perception in the intersubjective context. *Consciousness and cognition*, 17(2):535–543.

[Gallagher, 2011] Gallagher, S. (2011). *The Oxford handbook of the self.* Oxford University Press.

[Gallese and Goldman, 1998] Gallese, V. and Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in cognitive sciences*, 2(12):493–501.

[Gallese and Lakoff, 2005] Gallese, V. and Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive neuropsychology*, 22(3-4):455–479.

[García-Albea et al., 1996] García-Albea, J. E., del Viso Pabón, S., and Bernardos, M. L. S. (1996). *Test de Boston para el diagnóstico de la afasia: adaptación española*. Editorial Médica Panamericana.

[Gentile et al., 2013] Gentile, G., Guterstam, A., Brozzoli, C., and Ehrsson, H. H. (2013). Disintegration of multisensory signals from the real hand reduces default limb self-attribution: an fmri study. *Journal of Neuroscience*, 33(33):13350–13366.

[Gentile et al., 2011] Gentile, G., Petkova, V. I., and Ehrsson, H. H. (2011). Integration of visual and tactile signals from the hand in the human brain: an fmri study. *Journal of Neurophysiology*, 105(2):910–922.

[Gentilucci, 2003] Gentilucci, M. (2003). Grasp observation influences speech production. *European Journal of Neuroscience*, 17(1):179–184.

[Gerlach et al., 2002] Gerlach, C., Law, I., and Paulson, O. B. (2002). When action turns into words. activation of motor-based knowledge during categorization of manipulable objects. *Journal of cognitive neuroscience*, 14(8):1230–1239.

[Glenberg et al., 2009] Glenberg, A. M., Becker, R., Klötzer, S., Kolanko, L., Müller, S., and Rinck, M. (2009). Episodic affordances contribute to language comprehension. *Language and Cognition*, 1(1):113–135.

[Glenberg and Kaschak, 2002] Glenberg, A. M. and Kaschak, M. P. (2002). Grounding language in action. *Psychonomic bulletin & review*, 9(3):558–565.

[Glenberg and Robertson, 1999] Glenberg, A. M. and Robertson, D. A. (1999). Indexical understanding of instructions. *Discourse processes*, 28(1):1–26.

[Glenberg and Robertson, 2000] Glenberg, A. M. and Robertson, D. A. (2000). Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of memory and language*, 43(3):379–401.

[Glover and Dixon, 2002] Glover, S. and Dixon, P. (2002). Semantics affect the planning but not control of grasping. *Experimental Brain Research*, 146(3):383–387.

[González et al., 2006] González, J., Barros-Loscertales, A., Pulvermüller, F., Meseguer, V., Sanjuán, A., Belloch, V., and Ávila, C. (2006). Reading cinnamon activates olfactory brain regions. *Neuroimage*, 32(2):906–912.

[González-Franco et al., 2014] González-Franco, M., Peck, T. C., Rodríguez-Fornells, A., and Slater, M. (2014). A threat to a virtual hand elicits motor cortex activation. *Experimental brain research*, 232(3):875–887.

[Goodglass, 1980] Goodglass, H. (1980). Disorders of naming following brain injury. *American Scientist*, 68:647–655.

[Goodglass, 1993] Goodglass, H. (1993). *Understanding aphasia*. JSTOR.

[Gottlieb, 2007] Gottlieb, J. (2007). From thought to action: the parietal cortex as a bridge between perception, action, and cognition. *Neuron*, 53(1):9–16.

[Grafton et al., 1996] Grafton, S. T., Arbib, M. A., Fadiga, L., and Rizzolatti, G. (1996). Localization of grasp representations in humans by positron emission tomography. *Experimental brain research*, 112(1):103–111.

[Graziano and Botvinick, 2002] Graziano, M. S. and Botvinick, M. M. (2002). How the brain represents the body: insights from neurophysiology and psychology. *Common mechanisms in perception and action: Attention and performance XIX*, pages 136–157.

[Graziano et al., 2002] Graziano, M. S., Taylor, C. S., Moore, T., and Cooke, D. F. (2002). The cortical control of movement revisited. *Neuron*, 36(3):349–362.

[Grechuta et al., 2017a] Grechuta, K., Ballester, B. R., Munné, R. E., Bernal, T. U., Hervás, B. M., San Segundo, R., and Verschure, P. F. (2017a). The effects of silent visuomotor cueing on word retrieval in broca’s aphasies: A pilot study. In *2017 International Conference on Rehabilitation Robotics (ICORR)*, pages 193–199. IEEE.

[Grechuta et al., 2019a] Grechuta, K., Ballester, B. R., Munné, R. E., Bernal, T. U., Hervás, Begona Molina, B. M. F. P., San Segundo, R., and Verschure, P. F. (2019a). Augmented dyadic therapy boosts recovery of language function in patients with non-fluent aphasia:a randomised controlled trial. *Stroke*, In press.

[Grechuta et al., 2019b] Grechuta, K., Ballester, B. R., Munné, R. E., Bernal, T. U., Hervás, Begona Molina, B. M. F. P., San Segundo, R., and Verschure, P. F. (2019b). Automated multisensory cueing facilitates recovery in aphasia: a triple-blind clinical study. *Under review*.

[Grechuta et al., 2019c] Grechuta, K., De la Torre, J., Ballester, B. R., and Verschure, P. F. (2019c). Challenging the boundaries of the physical self: purely distal cues in the environment impact body ownership. *Royal Society Open Science*, In review.

[Grechuta et al., 2017b] Grechuta, K., Guga, J., Maffei, G., Rubio, B. B., and Verschure, P. F. (2017b). Visuotactile integration modulates motor performance in a perceptual decision-making task. *Scientific Reports*, 7.

[Grechuta et al., 2016a] Grechuta, K., Rubio, B., Duff, A., Duarte, E., and Verschure, P. (2016a). Intensive language-action therapy in virtual reality for a rehabilitation gaming system. *Journal of Pain Management*, 9:243.

[Grechuta et al., 2016b] Grechuta, K., Rubio, B., Duff, A., Duarte, E., and Verschure, P. (2016b). Intensive language-action therapy in virtual reality for a rehabilitation gaming system. *Journal of Pain Management*, 9:243.

[Grechuta et al., 2019d] Grechuta, K., Ulysse, L., Ballester, B. R., and Verschure, P. F. (2019d). Self beyond the body: action-driven and task-relevant purely distal cues modulate performance and body ownership. *Frontiers in Human Neuroscience*, 13:91.

[Griisser, 1995] Griisser, O.-J. (1995). On the history of the ideas of efference copy and reafference. In *Essays in the History of the Physiological Sciences: Proceedings of a Network Symposium of the European Association for the History of Medicine and Health Held at the University Louis Pasteur, Strasbourg, on March 26-27th, 1993*, volume 33, page 35. Rodopi.

[Grossman et al., 2008] Grossman, M., Anderson, C., Khan, A., Avants, B., Elman, L., and McCluskey, L. (2008). Impaired action knowledge in amyotrophic lateral sclerosis. *Neurology*, 71(18):1396–1401.

[Gunning et al., 2017] Gunning, D., Wenke, R., Ward, E. C., Chalk, S., Lawrie, M., Romano, M., Edwards, A., Hobson, T., and Cardell, E. (2017). Clinicians' perceptions of delivering new models of high intensity aphasia treatment. *Aphasiology*, 31(4):406–426.

[Guterstam et al., 2018] Guterstam, A., Collins, K. L., Cronin, J. A., Zeberg, H., Darvas, F., Weaver, K. E., Ojemann, J. G., and Ehrsson, H. H. (2018). Direct electrophysiological correlates of body ownership in human cerebral cortex. *Cerebral Cortex*, 29(3):1328–1341.

[Guterstam et al., 2019] Guterstam, A., Larsson, D. E., Zeberg, H., and Ehrsson, H. H. (2019). Multisensory correlations-not tactile expectations-determine the sense of body ownership. *PloS one*, 14(2):e0213265.

[Haans et al., 2008] Haans, A., IJsselsteijn, W. A., and de Kort, Y. A. (2008). The effect of similarities in skin texture and hand shape on perceived ownership of a fake limb. *Body Image*, 5(4):389–394.

[Hadidi et al., 2009] Hadidi, N., Treat-Jacobson, D. J., and Lindquist, R. (2009). Post-stroke depression and functional outcome: a critical review of literature. *Heart & Lung: The Journal of Acute and Critical Care*, 38(2):151–162.

[Haggard, 2005] Haggard, P. (2005). Conscious intention and motor cognition. *Trends in cognitive sciences*, 9(6):290–295.

[Haggard, 2017] Haggard, P. (2017). Sense of agency in the human brain. *Nature Reviews Neuroscience*, 18(4):196.

[Hägni et al., 2008] Hägni, K., Eng, K., Hepp-Reymond, M.-C., Holper, L., Keisker, B., Siekierka, E., and Kiper, D. C. (2008). Observing virtual arms that you imagine are yours increases the galvanic skin response to an unexpected threat. *PloS one*, 3(8):e3082.

[Hall et al., 2005] Hall, D. A., Fussell, C., and Summerfield, A. Q. (2005). Reading fluent speech from talking faces: typical brain networks and individual differences. *Journal of Cognitive Neuroscience*, 17(6):939–953.

[Halligan et al., 1995] Halligan, P. W., Marshall, J. C., and Wade, D. T. (1995). Unilateral somatoparaphrenia after right hemisphere stroke: a case description. *Cortex*, 31(1):173–182.

[Harnad, 1990] Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena*, 42(1-3):335–346.

[Hauk et al., 2004] Hauk, O., Johnsrude, I., and Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, 41(2):301–307.

[Head and Holmes, 1911] Head, H. and Holmes, G. (1911). Sensory disturbances from cerebral lesions. *Brain*, 34(2-3):102–254.

[Hecht and Reiner, 2009] Hecht, D. and Reiner, M. (2009). Sensory dominance in combinations of audio, visual and haptic stimuli. *Experimental brain research*, 193(2):307–314.

[Hilari et al., 2012] Hilari, K., Needle, J. J., and Harrison, K. L. (2012). What are the important factors in health-related quality of life for people with aphasia? a systematic review. *Archives of physical medicine and rehabilitation*, 93(1):S86–S95.

[Hoenig et al., 2008] Hoenig, K., Sim, E.-J., Bochev, V., Herrnberger, B., and Kiefer, M. (2008). Conceptual flexibility in the human brain: dynamic recruitment of semantic maps from visual, motor, and motion-related areas. *Journal of Cognitive Neuroscience*, 20(10):1799–1814.

[Holmes et al., 2004] Holmes, N. P., Calvert, G. A., and Spence, C. (2004). Extending or projecting peripersonal space with tools? multisensory interactions highlight only the distal and proximal ends of tools. *Neuroscience letters*, 372(1-2):62–67.

[Holmes and Spence, 2004] Holmes, N. P. and Spence, C. (2004). The body schema and multisensory representation (s) of peripersonal space. *Cognitive processing*, 5(2):94–105.

[Holmes and Spence, 2005] Holmes, N. P. and Spence, C. (2005). Multisensory integration: space, time and superadditivity. *Current Biology*, 15(18):R762–R764.

[Hommel, 2009] Hommel, B. (2009). Action control according to tec (theory of event coding). *Psychological Research PRPF*, 73(4):512–526.

[Howard and Papathanasiou, 2000] Howard, D. and Papathanasiou, I. (2000). Cognitive neuropsychology and aphasia therapy: The case of word retrieval. *Acquired neurogenic communication disorders: A clinical perspective*, pages 76–99.

[IJsselsteijn et al., 2006] IJsselsteijn, W. A., de Kort, Y. A. W., and Haans, A. (2006). Is this my hand i see before me? the rubber hand illusion in reality, virtual reality, and mixed reality. *Presence: Teleoperators and Virtual Environments*, 15(4):455–464.

[Iriki et al., 1996] Iriki, A., Tanaka, M., and Iwamura, Y. (1996). Coding of modified body schema during tool use by macaque postcentral neurones. *Neuroreport*, 7(14):2325–2330.

[Ishibashi et al., 2011] Ishibashi, R., Ralph, M. A. L., Saito, S., and Pobric, G. (2011). Different roles of lateral anterior temporal lobe and inferior parietal lobule in coding function and manipulation tool knowledge: evidence from an rtms study. *Neuropsychologia*, 49(5):1128–1135.

[Ito, 2008] Ito, M. (2008). Control of mental activities by internal models in the cerebellum. *Nature Reviews Neuroscience*, 9(4):304–313.

[Jack et al., 2001] Jack, D., Boian, R., Merians, A. S., Tremaine, M., Burdea, G. C., Adamovich, S. V., Recce, M., and Poizner, H. (2001). Virtual reality-enhanced stroke rehabilitation. *IEEE transactions on neural systems and rehabilitation engineering*, 9(3):308–318.

[Jackendoff and Jackendoff, 2002] Jackendoff, R. and Jackendoff, R. S. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford University Press, USA.

[James, 1890] James, W. (1890). The principles of psychology (vol. 1). *New York: Holt*, 474.

[James, 2013] James, W. (2013). *The principles of psychology*, volume 1. Read Books Ltd.

[Johansson, 2000] Johansson, B. B. (2000). Brain plasticity and stroke rehabilitation the willis lecture. *Stroke*, 31(1):223–230.

[Johnson and Haggard, 2005] Johnson, H. and Haggard, P. (2005). Motor awareness without perceptual awareness. *Neuropsychologia*, 43(2):227–237.

[Johnson et al., 2002] Johnson, H., van Beers, R. J., and Haggard, P. (2002). Action and awareness in pointing tasks. *Experimental Brain Research*, 146(4):451–459.

[Johnson-Laird, 1988] Johnson-Laird, P. N. (1988). *The computer and the mind: An introduction to cognitive science*. Harvard University Press.

[Jordan and Rumelhart, 1992] Jordan, M. I. and Rumelhart, D. E. (1992). Forward models: Supervised learning with a distal teacher. *Cognitive science*, 16(3):307–354.

[Kalckert et al., 2019] Kalckert, A., Bico, I., and Fong, J. X. (2019). Illusions with hands, but not with balloons–comparing ownership and referral of touch for a corporal and noncorporal object after visuotactile stimulation. *Perception*, page 0301006619839286.

[Kalckert and Ehrsson, 2012] Kalckert, A. and Ehrsson, H. H. (2012). Moving a rubber hand that feels like your own: a dissociation of ownership and agency. *Frontiers in human neuroscience*, 6:40.

[Kammers et al., 2009] Kammers, M. P., Longo, M. R., Tsakiris, M., Dijkerman, H. C., and Haggard, P. (2009). Specificity and coherence of body representations. *Perception*, 38(12):1804–1820.

[Karnath et al., 2005] Karnath, H.-O., Baier, B., and Nägele, T. (2005). Awareness of the functioning of one’s own limbs mediated by the insular cortex? *The Journal of Neuroscience*, 25(31):7134–7138.

[Kaschak and Glenberg, 2000] Kaschak, M. P. and Glenberg, A. M. (2000). Constructing meaning: The role of affordances and grammatical constructions in sentence comprehension. *Journal of memory and language*, 43(3):508–529.

[Katz et al., 2000] Katz, R. C., Hallowell, B., Code, C., Armstrong, E., Roberts, P., Pound, C., and Katz, L. (2000). A multinational comparison of aphasia management practices. *International Journal of Language & Communication Disorders*, 35(2):303–314.

[Katz and Wertz, 1997] Katz, R. C. and Wertz, R. T. (1997). The efficacy of computer-provided reading treatment for chronic aphasic adults. *Journal of Speech, Language, and Hearing Research*, 40(3):493–507.

[Kawato, 1999] Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current opinion in neurobiology*, 9(6):718–727.

[Keele, 1986] Keele, S. W. (1986). *Motor control.* John Wiley & Sons.

[Kendall et al., 2008] Kendall, D. L., Rosenbek, J. C., Heilman, K. M., Conway, T., Klenberg, K., Rothi, L. J. G., and Nadeau, S. E. (2008). Phoneme-based rehabilitation of anomia in aphasia. *Brain and language*, 105(1):1–17.

[Kersten et al., 2004] Kersten, D., Mamassian, P., and Yuille, A. (2004). Object perception as bayesian inference. *Annu. Rev. Psychol.*, 55:271–304.

[Kiefer and Pulvermüller, 2012] Kiefer, M. and Pulvermüller, F. (2012). Conceptual representations in mind and brain. *cortex*, 48:805–825.

[Kiefer et al., 2008] Kiefer, M., Sim, E.-J., Herrnberger, B., Grothe, J., and Hoenig, K. (2008). The sound of concepts: four markers for a link between auditory and conceptual brain systems. *Journal of Neuroscience*, 28(47):12224–12230.

[Kilteni et al., 2018] Kilteni, K., Andersson, B. J., Houborg, C., and Ehrsson, H. H. (2018). Motor imagery involves predicting the sensory consequences of the imagined movement. *Nature communications*, 9.

[Kilteni and Ehrsson, 2017] Kilteni, K. and Ehrsson, H. H. (2017). Body ownership determines the attenuation of self-generated tactile sensations. *Proceedings of the National Academy of Sciences*, 114(31):8426–8431.

[Kilteni et al., 2015] Kilteni, K., Maselli, A., Kording, K. P., and Slater, M. (2015). Over my fake body: body ownership illusions for studying the multisensory basis of own-body perception. *Frontiers in human neuroscience*, 9:141.

[Klatzky et al., 1989] Klatzky, R. L., Pellegrino, J. W., McCloskey, B. P., and Doherty, S. (1989). Can you squeeze a tomato? the role of motor representations in semantic sensibility judgments. *Journal of Memory and Language*, 28(1):56–77.

[Kleim and Jones, 2008] Kleim, J. A. and Jones, T. A. (2008). Principles of experience-dependent neural plasticity: implications for rehabilitation after brain damage. *Journal of speech, language, and hearing research*, 51(1):S225–S239.

[Kohler et al., 2002] Kohler, E., Keysers, C., Umilta, M. A., Fogassi, L., Gallese, V., and Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science*, 297(5582):846–848.

[Komitova et al., 2005] Komitova, M., Zhao, L. R., Gidö, G., Johansson, B. B., and Eriksson, P. (2005). Postischemic exercise attenuates whereas enriched environment has certain enhancing effects on lesion-induced subventricular zone activation in the adult rat. *European Journal of Neuroscience*, 21(9):2397–2405.

[Körding and Wolpert, 2004] Körding, K. P. and Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature*, 427(6971):244.

[Körding and Wolpert, 2006] Körding, K. P. and Wolpert, D. M. (2006). Bayesian decision theory in sensorimotor control. *Trends in cognitive sciences*, 10(7):319–326.

[Krakauer, 2009] Krakauer, J. W. (2009). Motor learning and consolidation: the case of visuomotor rotation. In *Progress in motor control*, pages 405–421. Springer.

[Krams et al., 1998] Krams, M., Rushworth, M., Deiber, M.-P., Frackowiak, R., and Passingham, R. (1998). The preparation, execution and suppression of copied movements in the human brain. *Experimental brain research*, 120(3):386–398.

[Kurland et al., 2018] Kurland, J., Liu, A., and Stokes, P. (2018). Effects of a tablet-based home practice program with telepractice on treatment outcomes in chronic aphasia. *Journal of Speech, Language, and Hearing Research*, 61(5):1140–1156.

[Kurland et al., 2012] Kurland, J., Pulvermüller, F., Silva, N., Burke, K., and Andrianopoulos, M. (2012). Constrained versus unconstrained intensive language therapy in two individuals with chronic, moderate-to-severe aphasia and apraxia of speech: behavioral and fmri outcomes. *American journal of speech-language pathology*, 21(2):S65–S87.

[Kurland et al., 2014] Kurland, J., Wilkins, A. R., and Stokes, P. (2014). ipractice: Piloting the effectiveness of a tablet-based home practice program in aphasia treatment. In *Seminars in speech and language*, volume 35, pages 051–064. Thieme Medical Publishers.

[Lakoff and Johnson, 1999] Lakoff, G. and Johnson, M. (1999). *Philosophy in the Flesh*, volume 4. New york: Basic books.

[Lakoff and Johnson, 2008] Lakoff, G. and Johnson, M. (2008). *Metaphors we live by*. University of Chicago press.

[Lavoie et al., 2017] Lavoie, M., Macoir, J., and Bier, N. (2017). Effectiveness of technologies in the treatment of post-stroke anomia: A systematic review. *Journal of communication disorders*, 65:43–53.

[Lee et al., 2010] Lee, J., Fowler, R., Rodney, D., Cherney, L., and Small, S. L. (2010). Imitate: An intensive computer-based treatment for aphasia based on action observation and imitation. *Aphasiology*, 24(4):449–465.

[Lenggenhager et al., 2007] Lenggenhager, B., Tadi, T., Metzinger, T., and Blanke, O. (2007). Video ergo sum: manipulating bodily self-consciousness. *Science*, 317(5841):1096–1099.

[Levelt, 1993] Levelt, W. J. (1993). *Speaking: From intention to articulation*, volume 1. MIT press.

[Levelt et al., 1999] Levelt, W. J., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and brain sciences*, 22(1):1–38.

[Lichteim, 1885] Lichteim, L. (1885). On aphasia. *Brain*, 7(4):433–484.

[Liepelt et al., 2017] Liepelt, R., Dolk, T., and Hommel, B. (2017). Self-perception beyond the body: the role of past agency. *Psychological research*, 81(3):549–559.

[Limanowski, 2014] Limanowski, J. (2014). What can body ownership illusions tell us about minimal phenomenal selfhood? *Frontiers in human neuroscience*, 8.

[Limanowski and Blankenburg, 2013] Limanowski, J. and Blankenburg, F. (2013). Minimal self-models and the free energy principle. *Frontiers in human neuroscience*, 7:547.

[Lloyd, 2007] Lloyd, D. M. (2007). Spatial limits on referred touch to an alien limb may reflect boundaries of visuo-tactile peripersonal space surrounding the hand. *Brain and cognition*, 64(1):104–109.

[Longo and Haggard, 2009] Longo, M. R. and Haggard, P. (2009). Sense of agency primes manual motor responses. *Perception*, 38(1):69–78.

[Longo et al., 2008] Longo, M. R., Schüür, F., Kammers, M. P., Tsakiris, M., and Haggard, P. (2008). What is embodiment? a psychometric approach. *Cognition*, 107(3):978–998.

[Lowell et al., 1995] Lowell, S., Beeson, P. M., and Holland, A. L. (1995). The efficacy of a semantic cueing procedure on naming performance of adults with aphasia. *American Journal of Speech-Language Pathology*, 4(4):109–114.

[Lyons et al., 2010] Lyons, I. M., Mattarella-Micke, A., Cieslak, M., Nusbaum, H. C., Small, S. L., and Beilock, S. L. (2010). The role of personal experience in the neural processing of action-related language. *Brain and language*, 112(3):214–222.

[Ma and Hommel, 2015a] Ma, K. and Hommel, B. (2015a). Body-ownership for actively operated non-corporeal objects. *Consciousness and Cognition*, 36:75–86.

[Ma and Hommel, 2015b] Ma, K. and Hommel, B. (2015b). The role of agency for perceived ownership in the virtual hand illusion. *Consciousness and cognition*, 36:277–288.

[MacKay and Crammond, 1987] MacKay, W. and Crammond, D. (1987). Neuronal correlates in posterior parietal lobe of the expectation of events. *Behavioural brain research*, 24(3):167–179.

[Maffei et al., 2017] Maffei, G., Herreros, I., Sanchez-Fibla, M., Friston, K. J., and Verschure, P. F. (2017). The perceptual shaping of anticipatory actions. *Proceedings of the Royal Society B: Biological Sciences*, 284(1869):20171780.

[Mahon and Caramazza, 2009] Mahon, B. Z. and Caramazza, A. (2009). Concepts and categories: A cognitive neuropsychological perspective. *Annual review of psychology*, 60:27–51.

[Maier et al., 2014] Maier, M., Ballester, B. R., Duarte, E., Duff, A., and Verschure, P. F. (2014). Social integration of stroke patients through the multiplayer rehabilitation gaming system. In *International Conference on Serious Games*, pages 100–114. Springer.

[Makin et al., 2008] Makin, T. R., Holmes, N. P., and Ehrsson, H. H. (2008). On the other hand: dummy hands and peripersonal space. *Behavioural brain research*, 191(1):1–10.

[Makin et al., 2007] Makin, T. R., Holmes, N. P., and Zohary, E. (2007). Is that near my hand? multisensory representation of peripersonal space in human intraparietal sulcus. *Journal of Neuroscience*, 27(4):731–740.

[Mamassian et al., 2002] Mamassian, P., Landy, M. S., and Maloney, L. T. (2002). Bayesian modelling of visual perception. *Probabilistic models of the brain: Perception and neural function*, pages 13–36.

[Maravita and Iriki, 2004] Maravita, A. and Iriki, A. (2004). Tools for the body (schema). *Trends in cognitive sciences*, 8(2):79–86.

[Maravita et al., 2003] Maravita, A., Spence, C., and Driver, J. (2003). Multisensory integration and the body schema: close to hand and within reach. *Current biology*, 13(13):R531–R539.

[Markie, 2004] Markie, P. (2004). Rationalism vs. empiricism.

[Marques, 2006] Marques, J. F. (2006). Specialization and semantic organization: Evidence for multiple semantics linked to sensory modalities. *Memory & Cognition*, 34(1):60–67.

[Marshall, 2008] Marshall, R. C. (2008). The impact of intensity of aphasia therapy on recovery. *Stroke*, 39(2):e48–e48.

[Martin, 2007] Martin, A. (2007). The representation of object concepts in the brain. *Annu. Rev. Psychol.*, 58:25–45.

[Martin and Chao, 2001] Martin, A. and Chao, L. L. (2001). Semantic memory and the brain: structure and processes. *Current opinion in neurobiology*, 11(2):194–201.

[Massaro and Palmer, 1998] Massaro, D. W. and Palmer, S. E. (1998). *Perceiving talking faces: From speech perception to a behavioral principle*, volume 1. Mit Press.

[Mayo et al., 2002] Mayo, N. E., Wood-Dauphinee, S., Côté, R., Durcan, L., and Carlton, J. (2002). Activity, participation, and quality of life 6 months poststroke. *Archives of physical medicine and rehabilitation*, 83(8):1035–1042.

[Mazzoni and Krakauer, 2006] Mazzoni, P. and Krakauer, J. W. (2006). An implicit plan overrides an explicit strategy during visuomotor adaptation. *Journal of neuroscience*, 26(14):3642–3645.

[McClelland and Rogers, 2003] McClelland, J. L. and Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nature reviews neuroscience*, 4(4):310.

[McNeil et al., 1990] McNeil, M., Tseng, C., and Lapointe, L. (1990). Aphasia and related neurogenic language disorders.

[Meinzer et al., 2007] Meinzer, M., Elbert, T., Djundja, D., Taub, E., and Rockstroh, B. (2007). Extending the constraint-induced movement therapy (cimt) approach to cognitive functions: Constraint-induced aphasia therapy (ciat) of chronic aphasia. *NeuroRehabilitation*, 22(4):311–318.

[Merians et al., 2002] Merians, A. S., Jack, D., Boian, R., Tremaine, M., Burdea, G. C., Adamovich, S. V., Recce, M., and Poizner, H. (2002). Virtual reality–augmented rehabilitation for patients following stroke. *Physical therapy*, 82(9):898–915.

[Merleau-Ponty, 1945] Merleau-Ponty, M. (1945). Phénoménologie de la perception.

[Meteyard et al., 2007] Meteyard, L., Bahrami, B., and Vigliocco, G. (2007). Motion detection and motion verbs: Language affects low-level visual perception. *Psychological Science*, 18(11):1007–1013.

[Meteyard et al., 2008] Meteyard, L., Zokaei, N., Bahrami, B., and Vigliocco, G. (2008). Visual motion interferes with lexical decision on motion words. *Current Biology*, 18(17):R732–R733.

[Miall and Wolpert, 1996] Miall, R. C. and Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural networks*, 9(8):1265–1279.

[Mihelj et al., 2014] Mihelj, M., Novak, D., and Beguš, S. (2014). *Virtual reality technology and applications*. Springer.

[Mohler et al., 2010] Mohler, B. J., Creem-Regehr, S. H., Thompson, W. B., and Bülthoff, H. H. (2010). The effect of viewing a self-avatar on distance judgments in an hmd-based virtual environment. *Presence: Teleoperators and Virtual Environments*, 19(3):230–242.

[mondiale de la santé and Organization, 2001] mondiale de la santé, O. and Organization, W. H. (2001). *International classification of functioning, disability and health: ICF.* World Health Organization.

[Morehead et al., 2017] Morehead, J. R., Taylor, J. A., Parvin, D. E., and Ivry, R. B. (2017). Characteristics of implicit sensorimotor adaptation revealed by task-irrelevant clamped feedback. *Journal of Cognitive Neuroscience*.

[Moseley et al., 2008] Moseley, G. L., Olthof, N., Venema, A., Don, S., Wijers, M., Gallace, A., and Spence, C. (2008). Psychologically induced cooling of a specific body part caused by the illusory ownership of an artificial counterpart. *Proceedings of the National Academy of Sciences*, 105(35):13169–13173.

[Murphy, 2006] Murphy, N. (2006). *Bodies and souls, or spirited bodies?*, volume 3. Cambridge University Press.

[Naeser et al., 2005] Naeser, M. A., Martin, P. I., Nicholas, M., Baker, E. H., Seekins, H., Kobayashi, M., Theoret, H., Fregni, F., Maria-Tormos, J., Kurland, J., et al. (2005). Improved picture naming in chronic aphasia after tms to part of right broca’s area: an open-protocol study. *Brain and language*, 93(1):95–105.

[Neville et al., 1998] Neville, H. J., Bavelier, D., Corina, D., Rauschecker, J., Karni, A., Lalwani, A., Braun, A., Clark, V., Jezzard, P., and Turner, R. (1998). Cerebral organization for language in deaf and hearing subjects: biological constraints and effects of experience. *Proceedings of the National Academy of Sciences*, 95(3):922–929.

[Newell and Simon, 1975] Newell, A. and Simon, H. A. (1975). Computer science as empirical inquiry: Symbols and search. *& &*.

[Newport et al., 2010] Newport, R., Pearce, R., and Preston, C. (2010). Fake hands in action: embodiment and control of supernumerary limbs. *Experimental brain research*, 204(3):385–395.

[Nickels, 2002] Nickels, L. (2002). Therapy for naming disorders. *Aphasiology*, 16:935–979.

[Nishitani and Hari, 2002] Nishitani, N. and Hari, R. (2002). Viewing lip forms: cortical dynamics. *Neuron*, 36(6):1211–1220.

[Nishitani et al., 2005] Nishitani, N., Schürmann, M., Amunts, K., and Hari, R. (2005). Broca’s region: from action to language. *Physiology*, 20(1):60–69.

[Noë, 2004] Noë, A. (2004). *Action in perception*. MIT press.

[Oldfield, 1971] Oldfield, R. C. (1971). The assessment and analysis of handedness: the edinburgh inventory. *Neuropsychologia*, 9(1):97–113.

[Oliveri et al., 2004] Oliveri, M., Finocchiaro, C., Shapiro, K., Gangitano, M., Caramazza, A., and Pascual-Leone, A. (2004). All talk and no action: a transcranial magnetic stimulation study of motor cortex activation during action word production. *Journal of cognitive neuroscience*, 16(3):374–381.

[O’Regan, 2011] O’Regan, J. K. (2011). *Why red doesn’t sound like a bell: Understanding the feel of consciousness*. Oxford University Press.

[O’Regan and Noë, 2001] O’Regan, J. K. and Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and brain sciences*, 24(5):939–973.

[Özyürek et al., 2007] Özyürek, A., Willems, R. M., Kita, S., and Hagoort, P. (2007). On-line integration of semantic information from speech and gesture: Insights from event-related brain potentials. *Journal of cognitive neuroscience*, 19(4):605–616.

[Page et al., 2012] Page, S. J., Fulk, G. D., and Boyne, P. (2012). Clinically important differences for the upper-extremity fugl-meyer scale in people with minimal to moderate impairment due to chronic stroke. *Physical therapy*, 92(6):791–798.

[Palmer et al., 2012] Palmer, R., Enderby, P., Cooper, C., Latimer, N., Julious, S., Paterson, G., Dimairo, M., Dixon, S., Mortley, J., Hilton, R., et al. (2012). Computer therapy compared with usual care for people with long-standing aphasia poststroke: a pilot randomized controlled trial. *Stroke*, 43(7):1904–1911.

[Palmer et al., 2013] Palmer, R., Enderby, P., and Paterson, G. (2013). Using computers to enable self-management of aphasia therapy exercises for word finding: the patient and carer perspective. *International journal of language & communication disorders*, 48(5):508–521.

[Pecher et al., 2003] Pecher, D., Zeelenberg, R., and Barsalou, L. W. (2003). Verifying different-modality properties for concepts produces switching costs. *Psychological science*, 14(2):119–124.

[Pecher and Zwaan, 2005] Pecher, D. and Zwaan, R. A. (2005). *Grounding cognition: The role of perception and action in memory, language, and thinking*. Cambridge University Press.

[Peck et al., 2013] Peck, T. C., Seinfeld, S., Aglioti, S. M., and Slater, M. (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and cognition*, 22(3):779–787.

[Peña-Casanova, 1990] Peña-Casanova, J. (1990). Test barcelona. *Ediciónes Masson, Barcelona*.

[Perrett et al., 1985] Perrett, D., Smith, P., Mistlin, A., Chitty, A., Head, A., Potter, D., Broennimann, R., Milner, A., and Jeeves, M. A. (1985). Visual analysis of body movements by neurones in the temporal cortex of the macaque monkey: a preliminary report. *Behavioural brain research*, 16(2-3):153–170.

[Petkova and Ehrsson, 2008] Petkova, V. I. and Ehrsson, H. H. (2008). If i were you: perceptual illusion of body swapping. *PloS one*, 3(12):e3832.

[Petrides et al., 2005] Petrides, M., Cadoret, G., and Mackey, S. (2005). Orofacial somatomotor responses in the macaque monkey homologue of broca’s area. *Nature*, 435(7046):1235.

[Pexman et al., 2007] Pexman, P. M., Hargreaves, I. S., Edwards, J. D., Henry, L. C., and Goodyear, B. G. (2007). Neural correlates of concreteness in semantic categorization. *Journal of Cognitive Neuroscience*, 19(8):1407–1419.

[Pia et al., 2004] Pia, L., Neppi-Modona, M., Ricci, R., and Berti, A. (2004). The anatomy of anosognosia for hemiplegia: a meta-analysis. *Cortex*, 40(2):367–377.

[Pobric et al., 2010] Pobric, G., Jefferies, E., and Ralph, M. A. L. (2010). Category-specific versus category-general semantic impairment induced by transcranial magnetic stimulation. *Current Biology*, 20(10):964–968.

[Prinz, 1997] Prinz, W. (1997). Perception and action planning. *European journal of cognitive psychology*, 9(2):129–154.

[Proske and Gandevia, 2012] Proske, U. and Gandevia, S. C. (2012). The proprioceptive senses: their roles in signaling body shape, body position and movement, and muscle force. *Physiological reviews*, 92(4):1651–1697.

[Pulvermüller, 2005] Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience*, 6(7):576.

[Pulvermüller and Berthier, 2008] Pulvermüller, F. and Berthier, M. L. (2008). Aphasia therapy on a neuroscience basis. *Aphasiology*, 22(6):563–599.

[Pulvermüller and Fadiga, 2010] Pulvermüller, F. and Fadiga, L. (2010). Active perception:sensorimotor circuits as a cortical basis for language. *Nature reviews neuroscience*, 11:351.

[Pulvermüller et al., 2001a] Pulvermüller, F., Härle, M., and Hummel, F. (2001a). Walking or talking?: Behavioral and neurophysiological correlates of action verb processing. *Brain and language*, 78(2):143–168.

[Pulvermüller et al., 2005] Pulvermüller, F., Hauk, O., Nikulin, V. V., and Ilmoniemi, R. J. (2005). Functional links between motor and language systems. *European Journal of Neuroscience*, 21(3):793–797.

[Pulvermüller et al., 2006] Pulvermüller, F., Huss, M., Kherif, F., del Prado Martin, F. M., Hauk, O., and Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences*, 103(20):7865–7870.

[Pulvermüller et al., 2016] Pulvermüller, F., Mohr, B., and Taub, E. (2016). Constraint-induced aphasia therapy: A neuroscience-centered translational method. In *Neurobiology of Language*, pages 1025–1034. Elsevier.

[Pulvermüller et al., 2001b] Pulvermüller, F., Neininger, B., Elbert, T., Mohr, B., Rockstroh, B., Koebbel, P., and Taub, E. (2001b). Constraint-induced therapy of chronic aphasia after stroke. *Stroke*, 32(7):1621–1626.

[Pulvermüller et al., 2003] Pulvermüller, F., Shtyrov, Y., and Ilmoniemi, R. (2003). Spatiotemporal dynamics of neural language processing: an meg study using minimum-norm current estimates. *Neuroimage*, 20(2):1020–1025.

[Putnam, 1963] Putnam, H. (1963). Brains and behavior.

[Pylyshyn, 1999] Pylyshyn, Z. (1999). Is vision continuous with cognition?: The case for cognitive impenetrability of visual perception. *Behavioral and brain sciences*, 22(3):341–365.

[Pylyshyn, 1984] Pylyshyn, Z. W. (1984). *Computation and cognition*. MIT press Cambridge, MA.

[Quillian, 1969] Quillian, M. R. (1969). The teachable language comprehender: A simulation program and theory of language. *Communications of the ACM*, 12(8):459–476.

[Ramachandran et al., 1998] Ramachandran, V. S., Blakeslee, S., and Sacks, O. W. (1998). *Phantoms in the brain: Probing the mysteries of the human mind*. William Morrow New York.

[Raymer et al., 2002] Raymer, A. M., Rothi, L., and Hillis, A. (2002). Clinical diagnosis and treatment of naming disorders. *The handbook of adult language disorders*, pages 163–182.

[Reisberg et al., 1987] Reisberg, D., McLean, J., and Goldfield, A. (1987). Easy to hear but hard to understand: a lip-reading advantage with intact auditory stimuli. b. dodd ve r. campbell,(ed.), hearing by eye: The psychology of lip-reading içinde (97-114).

[Revill et al., 2008] Revill, K. P., Aslin, R. N., Tanenhaus, M. K., and Bavelier, D. (2008). Neural correlates of partial lexical activation. *Proceedings of the National Academy of Sciences*, 105(35):13111–13115.

[Richards et al., 2008] Richards, L. G., Stewart, K. C., Woodbury, M. L., Senesac, C., and Cauraugh, J. H. (2008). Movement-dependent stroke recovery: a systematic review and meta-analysis of tms and fmri evidence. *Neuropsychologia*, 46(1):3–11.

[Rizzolatti and Arbib, 1998] Rizzolatti, G. and Arbib, M. A. (1998). Language within our grasp. *Trends in neurosciences*, 21(5):188–194.

[Rizzolatti and Craighero, 2004] Rizzolatti, G. and Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.*, 27:169–192.

[Rizzolatti et al., 1996] Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., and Fazio, F. (1996). Localization of grasp representations in humans by pet: 1. observation versus execution. *Experimental brain research*, 111(2):246–252.

[Rizzolatti et al., 1981a] Rizzolatti, G., Scandolara, C., Matelli, M., and Gentilucci, M. (1981a). Afferent properties of periarcuate neurons in macaque monkeys. i. somatosensory responses. *Behavioural brain research*, 2(2):125–146.

[Rizzolatti et al., 1981b] Rizzolatti, G., Scandolara, C., Matelli, M., and Gentilucci, M. (1981b). Afferent properties of periarcuate neurons in macaque monkeys. ii. visual responses. *Behavioural brain research*, 2(2):147–163.

[Rock and Victor, 1964] Rock, I. and Victor, J. (1964). Vision and touch: An experimentally created conflict between the two senses. *Science*, 143(3606):594–596.

[Rogers et al., 1999] Rogers, M. A., Alarcon, N. B., and Olswang, L. B. (1999). Aphasia management considered in the context of the world health organization model of disablements. *Physical Medicine and Rehabilitation Clinics*, 10(4):907–923.

[Rognini et al., 2013] Rognini, G., Sengül, A., Aspell, J., Salomon, R., Bleuler, H., and Blanke, O. (2013). Visuo-tactile integration and body ownership during self-generated action. *European Journal of Neuroscience*, 37(7):1120–1129.

[Romano et al., 2015] Romano, D., Caffa, E., Hernandez-Arieta, A., Brugger, P., and Maravita, A. (2015). The robot hand illusion: Inducing proprioceptive drift through visuo-motor congruency. *Neuropsychologia*, 70:414–420.

[Runes, 1960] Runes, D. D. (1960). Dictionary of philosophy (totowa, nj, littlefield, adams).

[Sacks and Sacks, 1998] Sacks, O. and Sacks, O. W. (1998). *A leg to stand on*. Simon and Schuster.

[Salomon et al., 2016] Salomon, R., Fernandez, N., van Elk, M., Vachicouras, N., Sabatier, F., Tychinskaya, A., Llobera, J., and Blanke, O. (2016). Changing motor perception by sensorimotor conflicts and body ownership. *Scientific reports*, 6.

[Samad et al., 2015] Samad, M., Chung, A. J., and Shams, L. (2015). Perception of body ownership is driven by bayesian sensory inference. *PloS one*, 10(2):e0117178.

[Sanchez-Vives and Slater, 2005] Sanchez-Vives, M. V. and Slater, M. (2005). From presence to consciousness through virtual reality. *Nature Reviews Neuroscience*, 6(4):332.

[Sanchez-Vives et al., 2010] Sanchez-Vives, M. V., Spanlang, B., Frisoli, A., Bergamasco, M., and Slater, M. (2010). Virtual hand illusion induced by visuomotor correlations. *PloS one*, 5(4):e10381.

[Saposnik et al., 2010] Saposnik, G., Teasell, R., Mamdani, M., Hall, J., McIlroy, W., Cheung, D., Thorpe, K. E., Cohen, L. G., Bayley, M., Group, S. O. R. C. S. W., et al. (2010). Effectiveness of virtual reality using wii gaming technology in stroke rehabilitation a pilot randomized clinical trial and proof of principle. *Stroke*, 41(7):1477–1484.

[Sato and Yasuda, 2005] Sato, A. and Yasuda, A. (2005). Illusion of sense of self-agency: discrepancy between the predicted and actual sensory consequences of actions modulates the sense of self-agency, but not the sense of self-ownership. *Cognition*, 94(3):241–255.

[Schwarz et al., 2018] Schwarz, K. A., Burger, S., Dignath, D., Kunde, W., and Pfister, R. (2018). Action-effect binding and agency. *Consciousness and cognition*, 65:304–309.

[Schwoebel et al., 2001] Schwoebel, J., Friedman, R., Duda, N., and Coslett, H. B. (2001). Pain and the body schema. *Brain*, 124(10):2098–2104.

[Searle, 1980] Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and brain sciences*, 3(3):417–424.

[Serino et al., 2015] Serino, A., Canzoneri, E., Marzolla, M., Di Pellegrino, G., and Magosso, E. (2015). Extending peripersonal space representation without tool-use: evidence from a combined behavioral-computational approach. *Frontiers in Behavioral Neuroscience*, 9:4.

[Seth, 2013] Seth, A. K. (2013). Interoceptive inference, emotion, and the embodied self. *Trends in cognitive sciences*, 17(11):565–573.

[Sforza et al., 2010] Sforza, A., Bufalari, I., Haggard, P., and Aglioti, S. M. (2010). My face in yours: Visuo-tactile facial stimulation influences sense of identity. *Social neuroscience*, 5(2):148–162.

[Shadmehr et al., 1994] Shadmehr, R., Mussa-Ivaldi, F. A., et al. (1994). Adaptive representation of dynamics during learning of a motor task. *Journal of Neuroscience*, 14(5):3208–3224.

[Shadmehr et al., 2010] Shadmehr, R., Smith, M. A., and Krakauer, J. W. (2010). Error correction, sensory prediction, and adaptation in motor control. *Annual review of neuroscience*, 33:89–108.

[Shergill et al., 2003] Shergill, S. S., Bays, P. M., Frith, C. D., and Wolpert, D. M. (2003). Two eyes for an eye: the neuroscience of force escalation. *Science*, 301(5630):187–187.

[Shibuya et al., 2018] Shibuya, S., Unenaka, S., and Ohki, Y. (2018). The relationship between the virtual hand illusion and motor performance. *Frontiers in psychology*, 9:2242.

[Shin et al., 2010] Shin, Y. K., Proctor, R. W., and Capaldi, E. J. (2010). A review of contemporary ideomotor theory. *Psychological bulletin*, 136(6):943.

[Short and Ward, 2009] Short, F. and Ward, R. (2009). Virtual limbs and body space: Critical features for the distinction between body space and near-body space. *Journal of Experimental Psychology: Human Perception and Performance*, 35(4):1092.

[Simmons, 2003] Simmons, A. (2003). Descartes on the cognitive structure of sensory experience. *Philosophy and Phenomenological Research*, 67(3):549–579.

[Skipper et al., 2007] Skipper, J. I., Goldin-Meadow, S., Nusbaum, H. C., and Small, S. L. (2007). Speech-associated gestures, broca's area, and the human mirror system. *Brain and language*, 101(3):260–277.

[Slater, 2009] Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535):3549–3557.

[Slater et al., 2008] Slater, M., Pérez Marcos, D., Ehrsson, H., and Sanchez-Vives, M. V. (2008). Towards a digital body: the virtual arm illusion. *Frontiers in human neuroscience*, 2:6.

[Smit et al., 2018] Smit, M., Brummelman, J., Keizer, A., van der Smagt, M., Dijkerman, H., and van der Ham, I. (2018). Body ownership and the absence of touch: approaching the rubber hand inside and outside peri-hand space. *Experimental brain research*, pages 1–15.

[Smith, 2004] Smith, A. D. (2004). *The problem of perception*. Motilal Banarsidass Publishe.

[Snodgrass and Vanderwart, 1980] Snodgrass, J. G. and Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of experimental psychology: Human learning and memory*, 6(2):174.

[Sober and Sabes, 2003] Sober, S. J. and Sabes, P. N. (2003). Multisensory integration during motor planning. *Journal of Neuroscience*, 23(18):6982–6992.

[Sommer and Wurtz, 2008] Sommer, M. A. and Wurtz, R. H. (2008). Brain circuits for the internal monitoring of movements. *Annu. Rev. Neurosci.*, 31:317–338.

[Spence et al., 2001] Spence, C., Nicholls, M. E., and Driver, J. (2001). The cost of expecting events in the wrong sensory modality. *Perception & Psychophysics*, 63(2):330–336.

[Sperry, 1950] Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of comparative and physiological psychology*, 43(6):482.

[Stahl et al., 2016] Stahl, B., Mohr, B., Dreyer, F. R., Lucchese, G., and Pulvermüller, F. (2016). Using language for social interaction: communication mechanisms promote recovery from chronic non-fluent aphasia. *Cortex*, 85:90–99.

[Stanfield and Zwaan, 2001] Stanfield, R. A. and Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological science*, 12(2):153–156.

[Stark and Warburton, 2016] Stark, B. C. and Warburton, E. A. (2016). Improved language in chronic aphasia after self-delivered ipad speech therapy. *Neuropsychological rehabilitation*, pages 1–14.

[Stein, 1998] Stein, B. E. (1998). Neural mechanisms for synthesizing sensory information and producing adaptive behaviors. *Experimental Brain Research*, 123(1):124–135.

[Stein and Stanford, 2008] Stein, B. E. and Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nature Reviews Neuroscience*, 9(4):255–266.

[Stock and Stock, 2004] Stock, A. and Stock, C. (2004). A short history of ideo-motor action. *Psychological research*, 68(2-3):176–188.

[Suzuki et al., 2013] Suzuki, K., Garfinkel, S. N., Critchley, H. D., and Seth, A. K. (2013). Multisensory integration across exteroceptive and interoceptive domains modulates self-experience in the rubber-hand illusion. *Neuropsychologia*, 51(13):2909–2917.

[Taimela, 1991] Taimela, S. (1991). Factors affecting reaction-time testing and the interpretation of results. *Perceptual and Motor skills*, 73(3 suppl):1195–1202.

[Taub et al., 2006] Taub, E., Uswatte, G., Mark, V., and Morris, D. (2006). The learl1ed nonuse phenomenon: implications for rehabilitation. *Eura Medicophys*, 42:241–55.

[Thompson et al., 2008] Thompson, C. K., Worrall, L., and Martin, N. (2008). Approaches to aphasia treatment. *Aphasia rehabilitation. The impairment and its consequences*, pages 3–24.

[Tsakiris, 2010] Tsakiris, M. (2010). My body in the brain: a neurocognitive model of body-ownership. *Neuropsychologia*, 48(3):703–712.

[Tsakiris et al., 2010] Tsakiris, M., Carpenter, L., James, D., and Fotopoulou, A. (2010). Hands only illusion: multisensory integration elicits sense of ownership for body parts but not for non-corporeal objects. *Experimental Brain Research*, 204(3):343–352.

[Tsakiris et al., 2008] Tsakiris, M., Costantini, M., and Haggard, P. (2008). The role of the right temporo-parietal junction in maintaining a coherent sense of one's body. *Neuropsychologia*, 46(12):3014–3018.

[Tsakiris and Haggard, 2005] Tsakiris, M. and Haggard, P. (2005). The rubber hand illusion revisited: visuotactile integration and self-attribution. *Journal of Experimental Psychology: Human Perception and Performance*, 31(1):80.

[Tsakiris et al., 2005] Tsakiris, M., Haggard, P., Franck, N., Mainy, N., and Sirigu, A. (2005). A specific role for efferent information in self-recognition. *Cognition*, 96(3):215–231.

[Tsakiris et al., 2006a] Tsakiris, M., Hesse, M. D., Boy, C., Haggard, P., and Fink, G. R. (2006a). Neural signatures of body ownership: a sensory network for bodily self-consciousness. *Cerebral cortex*, 17(10):2235–2244.

[Tsakiris et al., 2007] Tsakiris, M., Hesse, M. D., Boy, C., Haggard, P., and Fink, G. R. (2007). Neural signatures of body ownership: a sensory network for bodily self-consciousness. *Cerebral cortex*, 17(10):2235–2244.

[Tsakiris et al., 2006b] Tsakiris, M., Prabhu, G., and Haggard, P. (2006b). Having a body versus moving your body: How agency structures body-ownership. *Consciousness and cognition*, 15(2):423–432.

[Tseng et al., 2007] Tseng, Y.-w., Diedrichsen, J., Krakauer, J. W., Shadmehr, R., and Bastian, A. J. (2007). Sensory prediction errors drive cerebellum-dependent adaptation of reaching. *Journal of neurophysiology*, 98(1):54–62.

[Tucker and Ellis, 2001] Tucker, M. and Ellis, R. (2001). The potentiation of grasp types during visual object categorization. *Visual cognition*, 8(6):769–800.

[Tyler and Moss, 2001] Tyler, L. K. and Moss, H. E. (2001). Towards a distributed account of conceptual knowledge. *Trends in cognitive sciences*, 5(6):244–252.

[Vallar and Ronchi, 2009] Vallar, G. and Ronchi, R. (2009). Somatoparaphrenia: a body delusion. a review of the neuropsychological literature. *Experimental Brain Research*, 192(3):533–551.

[Van Dam and Stephens, 2018] Van Dam, L. C. and Stephens, J. R. (2018). Effects of prolonged exposure to feedback delay on the qualitative subjective experience of virtual reality. *PloS one*, 13(10):e0205145.

[van de Sandt-Koenderman, 2011] van de Sandt-Koenderman, W. M. E. (2011). Aphasia rehabilitation and the role of computer technology: Can we keep up with modern times? *International Journal of Speech-Language Pathology*, 13(1):21–27.

[Van Den Bos and Jeannerod, 2002] Van Den Bos, E. and Jeannerod, M. (2002). Sense of body and sense of action both contribute to self-recognition. *Cognition*, 85(2):177–187.

[van Ee et al., 2009] van Ee, R., van Boxtel, J. J., Parker, A. L., and Alais, D. (2009). Multisensory congruency as a mechanism for attentional control over perceptual selection. *The Journal of Neuroscience*, 29(37):11641–11649.

[Van Gelder and Port, 1995] Van Gelder, T. and Port, R. F. (1995). It’s about time: An overview of the dynamical approach to cognition. *Mind as motion: Explorations in the dynamics of cognition*, 1:43.

[Venna et al., 2014] Venna, V., Xu, Y., Doran, S., Patrizz, A., and McCullough, L. (2014). Social interaction plays a critical role in neurogenesis and recovery after stroke. *Translational psychiatry*, 4(1):e351.

[Vermeulen et al., 2007] Vermeulen, N., Niedenthal, P. M., and Luminet, O. (2007). Switching between sensory and affective systems incurs processing costs. *Cognitive Science*, 31(1):183–192.

[Verschure, 2011] Verschure, P. F. (2011). Neuroscience, virtual reality and neurorehabilitation: brain repair as a validation of brain theory. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 2254–2257. IEEE.

[Verschure and Althaus, 2003] Verschure, P. F. and Althaus, P. (2003). A real-world rational agent: Unifying old and new ai. *Cognitive science*, 27(4):561–590.

[Verschure et al., 2003] Verschure, P. F., Voegtlin, T., and Douglas, R. J. (2003). Environmentally mediated synergy between perception and behaviour in mobile robots. *Nature*, 425(6958):620.

[Vogeley et al., 2004] Vogeley, K., May, M., Ritzl, A., Falkai, P., Zilles, K., and Fink, G. R. (2004). Neural correlates of first-person perspective as one constituent of human self-consciousness. *Journal of cognitive neuroscience*, 16(5):817–827.

[Walsh et al., 2011] Walsh, L. D., Moseley, G. L., Taylor, J. L., and Gandevia, S. C. (2011). Proprioceptive signals contribute to the sense of body ownership. *The Journal of physiology*, 589(12):3009–3021.

[Warrington and McCarthy, 1987] Warrington, E. K. and McCarthy, R. A. (1987). Categories of knowledge: Further fractionations and an attempted integration. *Brain*, 110(5):1273–1296.

[Watkins and Paus, 2004] Watkins, K. and Paus, T. (2004). Modulation of motor excitability during speech perception: the role of broca’s area. *Journal of cognitive neuroscience*, 16(6):978–987.

[Wertheimer, 2012] Wertheimer, M. (2012). *A brief history of psychology*. Psychology Press.

[Whyte and Mulsant, 2002] Whyte, E. M. and Mulsant, B. H. (2002). Post stroke depression: epidemiology, pathophysiology, and biological treatment. *Biological psychiatry*, 52(3):253–264.

[Willems et al., 2010] Willems, R. M., Hagoort, P., and Casasanto, D. (2010). Body-specific representations of action verbs: Neural evidence from right-and left-handers. *Psychological Science*, 21(1):67–74.

[Wilson and Foglia, 2011] Wilson, R. A. and Foglia, L. (2011). Embodied cognition.

[Wilson-Mendenhall et al., 2013] Wilson-Mendenhall, C. D., Simmons, W. K., Martin, A., and Barsalou, L. W. (2013). Contextual processing of abstract concepts reveals neural representations of nonlinguistic semantic content. *Journal of cognitive neuroscience*, 25(6):920–935.

[Witmer, 1998] Witmer, B. G. (1998). Measuring presence. pages 225–240.

[Witmer and Singer, 1998] Witmer, B. G. and Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence*, 7(3):225–240.

[Witt, 1999] Witt, S. M. (1999). *Use of speech recognition in computer-assisted language learning*. PhD thesis, University of Cambridge Cambridge, United Kingdom.

[Wolpert et al., 2011] Wolpert, D. M., Diedrichsen, J., and Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nature reviews. Neuroscience*, 12(12):739.

[Wolpert and Flanagan, 2001] Wolpert, D. M. and Flanagan, J. R. (2001). Motor prediction. *Current biology*, 11(18):R729–R732.

[Wolpert et al., 1995] Wolpert, D. M., Ghahramani, Z., and Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science*, 269(5232):1880–1882.

[Wolpert and Kawato, 1998] Wolpert, D. M. and Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural networks*, 11(7):1317–1329.

[Woodgate et al., 2015] Woodgate, P. J., Strauss, S., Sami, S. A., and Heinke, D. (2015). Motor cortex guides selection of predictable movement targets. *Behavioural brain research*, 287:238–246.

[Wyss et al., 2004] Wyss, R., König, P., and Verschure, P. F. (2004). Involving the motor system in decision making. *Proceedings of the Royal Society of London B: Biological Sciences*, 271(Suppl 3):S50–S52.

[Xerri et al., 1998] Xerri, C., Merzenich, M. M., Peterson, B. E., and Jenkins, W. (1998). Plasticity of primary somatosensory cortex paralleling sensorimotor skill recovery from stroke in adult monkeys. *Journal of neurophysiology*, 79(4):2119–2148.

[Yamamoto and Kitazawa, 2001] Yamamoto, S. and Kitazawa, S. (2001). Reversal of subjective temporal order due to arm crossing. *Nature neuroscience*, 4(7):759.

[Zahavi and Parnas, 1998] Zahavi, D. and Parnas, J. (1998). Phenomenal consciousness and self-awareness: A phenomenological critique of representational theory. *Journal of Consciousness Studies*, 5(5-6):687–705.

[Zhang and Rowe, 2016] Zhang, J. and Rowe, J. B. (2016). Dissociable mechanisms of speed-accuracy tradeoff during visual perceptual learning are revealed by a hierarchical drift-diffusion model. *Toward a Unified View of the Speed-Accuracy Trade-Off: Behaviour, Neurophysiology and Modelling*, page 71.

[Zheng et al., 2016] Zheng, C., Lynch, L., and Taylor, N. (2016). Effect of computer therapy in aphasia: a systematic review. *Aphasiology*, 30(2-3):211–244.

[Zwaan et al., 2004] Zwaan, R. A., Madden, C. J., Yaxley, R. H., and Aveyard, M. E. (2004). Moving words: Dynamic representations in language comprehension. *Cognitive Science*, 28(4):611–619.

[Zwaan et al., 2002] Zwaan, R. A., Stanfield, R. A., and Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological science*, 13(2):168–171.