ECOLE DOCTORALE COGNITON, LANGAGE, EDUCATION (356)
LABORATOIRE PAROLE ET LANGAGE

THÈSE
en vue de l'obtention du grade de
DOCTEUR de L'UNIVERSITÉ D'AIX-MARSEILLE
Discipline : Sciences du Langage

Présentée par
ANA ZAPPA

le 10 décembre 2020

EMBODIED SEMANTICS PUT TO THE TEST

ELECTROPHYSIOLOGICAL EVIDENCE FROM VIRTUAL REALITY AND CLASSICAL ENVIRONMENTS

membres du jury:
Daniel MESTRE, Directeur de recherche, CNRS, *président*
Yann COELLO, Professeur, Université de Lille, *rapporteur*
Véronique BOULENGER, Chargée de rercherche, CNRS, *examinatrice*
Alice FOUCART, Professeure, Universidad Nebrija, *rapportrice*
Cheryl FRENCK-MESTRE, Directrice de recherche, CNRS, *directrice de thèse*

Abstract

While neuroimaging and behavioral studies have shown that sensory-motor systems are recruited during semantic processing, how and when this occurs has not yet been clearly established. Our purpose was to observe the different contexts in which motor activation can contribute to language comprehension and learning, using interactive and ecologically valid environments. In our first study, novice learners acquired a reduced second-language (L2) lexicon through interactive computer-games. Behavioral and electroencephalography (EEG) results indexed rapid L2 word learning. Interestingly, even-related potential (ERP) results revealed a gender congruency effect such that only words that had the same grammatical gender across participants' L1 and L2 gave rise to an N400 effect for match vs mismatch auditory word and image pairs, indicating that these words were better encoded. In a second study, we used an action-sentence compatibility effect (ACE) paradigm to evaluate how motor preparation affects language processing. ERP results showed greater N400 amplitude for congruent compared to incongruent action sentence trials, suggesting that compatibility between motor and language processes produced interference. In studies 3 and 4, we combined virtual reality (VR) and EEG to investigate interactions between language processing and motor activation. In the first of these studies, participants heard action verbs in their native language and performed varied actions on a virtual object in a Cave automatic virtual environment (CAVE) during a Go-Nogo task. Time-frequency analysis showed motor activation for both Go and Nogo conditions during action verb processing and prior to movement proper. In addition, greater motor activation for Go versus Nogo trials. Our final (projected) study is a registered report that aims to determine the neural correlates of embodied L2 learning by having participants encode auditory action verbs using an interactive virtual reality head-mounted display system and specific real-life actions on a virtual object. Using behavioral and EEG measures in a pre-post training design, this condition will be compared to a control condition in which participants will simply point to the virtual object.

The body of the work reported in this dissertation represents a significant step towards better understanding the subtle relationship between motor and semantic processes. By making use of new technologies that allow for manipulating and controlling the environment, our work opens up fresh perspectives for taking into account the contextual nature of how we learn and understand language.

Résumé

Bien que la neuro-imagerie et les études comportementales aient montré un recrutement des systèmes sensorimoteurs lors du traitement sémantique, le rôle et le décours temporel de ces activations doivent encore être précisés. Notre objectif était donc d'observer les différents contextes dans lesquels l'activation motrice peut contribuer à la compréhension et à l'apprentissage du langage, en utilisant des environnements interactifs et écologiques. Dans une première étude, des apprenants novices ont acquis un vocabulaire restreint dans une deuxième langue (L2) au travers de jeux interactifs sur ordinateur. Les résultats comportementaux et en électroencéphalographie (EEG) ont indiqué un apprentissage rapide de mots en L2. Par contre, les potentiels évoqués ont révélé un effet de congruence de genre grammatical : seuls les mots ayant le même genre dans la première et la deuxième langue ont produit un effet N400 (réponse à la congruence/incongruence de paires de mots auditifs et d'images), indiquant que ceux-ci ont été mieux intégrés. Dans une deuxième étude, nous avons utilisé un paradigme de compatibilité entre phrases d'action et actions réalisées par le participant (Action-sentence compatibility effect, ou ACE) pour mesurer les effets de la préparation motrice sur le traitement du langage. Les résultats en EEG ont montré une plus grande amplitude de la N400 pour des essais congruents par rapport aux essais incongruents, ce qui suggère une interférence provoquée par la compatibilité entre les processus moteurs et langagiers. Dans les études 3 et 4, nous avons combiné la réalité virtuelle (RV) et l'EEG pour analyser des interactions entre le traitement du langage et l'activation motrice. Dans l'étude 3, les participants ont entendu des verbes d'action dans leur langue maternelle et ont effectué diverses actions sur un objet virtuel dans un environnement virtuel hautement immersif, le Cave automatic virtual environment (CAVE), pendant une tâche de Go-Nogo. Des analyses en temps-fréquence du signal EEG ont montré des activations motrices dans les deux conditions, avant tout mouvement physique. Cette activation était plus importante pour les essais « Go » que pour les essais « Nogo ». Notre dernière étude (prévue) est un registered report qui vise à déterminer les corrélats neuronaux de l'apprentissage incarné d'une deuxième langue. Les participants encoderont des verbes d'action auditifs à l'aide d'un casque de réalité virtuelle interactive en effectuant des actions spécifiques pour manipuler un objet virtuel. Dans un design pré-/post-entraînement, l'apprentissage sera évalué avec des mesures EEG et comportementales et comparé à l'apprentissage dans une condition où les participants pointent vers l'objet.

Le travail présenté dans cette thèse constitue un pas important vers une meilleure compréhension de la relation subtile entre les processus moteurs et sémantiques. En tirant profit des nouvelles technologies qui permettent de manipuler et contrôler l'environnement, notre travail ouvre de nouvelles perspectives pour prendre en compte la nature contextuelle de notre apprentissage et de notre compréhension du langage.

Dedication

Pra Juju, claro.

Acknowledgements

Allocation de Bourse de thèse MENRT (AMU, Ecole Doctorale 356)

Research funding from the ILCB

Thank you, first and foremost, Cheryl... I could not have hoped for a better mentor. You embody the most admirable aspects of research and your passion for it is contagious. Thank you for your patience, your fastidiousness, your curiosity, your extremely hard work and your humanity. You must know how much I appreciate your having believed in me but if there's one place to put it down in words (or attempt to), it's here.

Je remercie Yann Coello et Alice Foucart d'avoir accepté d'être rapporteurs de ma thèse et Véronique Boulenger et Daniel Mestre d'avoir accepté de participer à mon jury.

Merci à l'ensemble des membres du LPL pour vos enseignements, pour votre "input", pour vos encouragements, pour les collaborations et pour cette belle énergie dans le jardin à midi.

Merci Stéphanie, Nadera, Cathy, Nadia, Antonio, Sebastien et Cyril pour votre travail constant et pour toute votre aide au quotidien.

Je remercie tous mes participants, qui ont rendu cette aventure encore plus enrichissante.

Agradeço aos meus participantes, que tanto enriqueceram este meu trabalho.

Dee, you are the most brilliant (and not useless like a light in the desert!) and patient person I know; thank you for the hours and hours of teaching and work... and for your creativity and investment in these studies.

Daniel, merci pour ton soutien constant, tes idées innovantes et ta presence apaisante.

Merci, Jean-Marie, pour ton travail méticuleux, ton ouverture d'esprit et de ton dynamisme.

Céline, un énorme merci; sans toi je n'aurais jamais fini la Licence, ni le Master... quelle chance d'avoir été assise à coté de la fille la plus généreuse, intélligente et drôle de la classe.

Merci, Cristel Portes, de m'avoir mise sur la bonne piste et de m'avoir soutenue pendant toutes ces années.

Merci Noël Nguyen pour le soutien essentiel au tout début.

Et un grand merci également à...

Raphaël pour les conversations incarnées et inspiratrices, les collaborations, la relecture et les conseils précieux.

Daria et Fatima pour votre aide indispensable lors des passassions…

Mes camarades doctorants et jeunes docteurs qui ont rendu ce voyage académique encore plus riche et drôle… vous m'inspirez !

Mes amis Marseillais de m'avoir supportée, soutenue et encouragée… et de m'avoir permis de souffler de temps en temps.

My far-away friends, who remain present, day in and day out.

Minha família e amigos-família no Brasil, não sou ninguém sem vocês.

Ma très belle-famille, Pierre, Anne et Annie - chercheuse inspiratrice ! - de m'avoir soutenue.

MEB for teaching me to ask questions, for being the most encouraging and supportive mother in the world (I've done the stats), for your meticulous proof-reading, and for putting up with me, unconditionally...

Pai, pelo amor e pelos neurônios!

Vovó e Nando, sempre presentes...

Jean-Luc, de m'avoir toujours encouragée et, je sais, je sais, supportée pendant ce travail de longue halaine… ça n'aurait jamais été possible sans toi…

Juliette, que me inspira, me alegra e me enche de energia, com vontade de fazer cada vez mais… um dia você vai entender que fiz essa tese, também, por você…

Je sors de cette expérience with a great sense of accomplishment, mas, mais que tudo, com a certeza de ter beaucoup de chance de vous avoir in my life.

Contents

1 The theoretical context 1

1.1 General Introduction 1

1.1.1 The symbol grounding problem and embodied semantics 2

1.1.2 Semantic representations 2

1.1.3 The complexity of putting embodied semantics to the test 4

1.1.4 Contextual language processing and embodiment 6

1.2 Electroencephalography to measure language processing 7

1.2.1 Event-related potentials (ERPs) 7

1.2.2 Time-Frequency ERSP to measure motor activation during language processing 8

1.3 Using the ACE to measure interactions between motor and linguistic processes 9

1.3.1 The neural correlates of the ACE 10

1.3.2 Further exploring the neural correlates of the ACE using EEG 11

1.4 Adult L2 learning 12

1.4.1 Games for interactive L2 learning 13

1.4.2 Embodied semantics and language learning: behavioral evidence . . . 14

1.4.3 Learning studies to explore the neural correlates of embodied encoding 15

1.5 Virtual Reality 16

1.5.1 VR: a naturalistic environment in which to measure language processing 16

1.5.2 Combining VR and EEG in language processing and learning studies . 18

1.6 Completed Studies 19

1.6.1 Cross-linguistic gender congruency effects during lexical access in novice L2 learners: evidence from EEG 19

1.6.2 What role does motor activation play in action language processing? An EEG study 20

1.6.3 Motor Resonance during linguistic processing as shown by EEG in a naturalistic VR environment . . . 21
1.6.4 The neural correlates of embodied L2 learning. Does embodied L2 verb learning affect representation and retention? . . . 21

2 Cross-linguistic gender congruency effects during lexical access in novice L2 learners: evidence from EEG 23
2.1 Introduction . . . 25
2.2 Methods . . . 30
2.2.1 Participants . . . 30
2.2.2 Materials . . . 30
2.2.3 Procedure . . . 32
2.2.4 EEG Match-Mismatch task . . . 33
2.2.5 ERP Data acquisition and analysis . . . 33
2.3 Results . . . 34
2.3.1 Behavioral results . . . 34
2.3.2 ERP analysis . . . 35
2.4 Tables . . . 41
2.5 Discussion . . . 49

3 What role does motor activation play in action language processing? An EEG study 61
3.1 Introduction . . . 62
3.2 Methods . . . 68
3.2.1 Participants . . . 68
3.2.2 Materials . . . 68
3.2.3 Procedure . . . 70
3.2.4 ERP Data Acquisition and Processing . . . 70
3.2.5 Event-related Spectral Perturbation (ERSP) . . . 71
3.3 Results . . . 71
3.3.1 Behavioral measures . . . 71
3.3.2 ERP analysis . . . 73
3.3.3 ERSP analysis . . . 76
3.4 Discussion . . . 81
3.4.1 ERPs . . . 81

3.4.2 Behavioral effects . . . 83
3.4.3 Time-frequency effects . . . 84
3.4.4 Conclusion . . . 85
3.5 Annex . . . 86

4 Motor Resonance during linguistic processing as shown by EEG in a naturalistic VR environment 95

5 Brain Waves in Virtual Reality 111
5.1 Introduction . . . 112
5.2 Methods . . . 113
5.2.1 Stimulation system . . . 113
5.2.2 Acquisition system: Electroencephalograpy . . . 117
5.2.3 Synchronization of stimulation and acquisition systems . . . 117
5.3 Conclusion . . . 120

6 The neural correlates of embodied L2 learning 123
6.1 Introduction . . . 124
6.1.1 Hypotheses . . . 130
6.2 Methods . . . 132
6.2.1 Ethics . . . 132
6.2.2 Statistical Power Analysis . . . 132
6.2.3 Participants . . . 132
6.2.4 Stimuli . . . 133
6.2.5 Learning apparatus . . . 134
6.2.6 Software . . . 134
6.2.7 General Procedure . . . 134
6.2.8 Learning Procedure . . . 134
6.2.9 EEG procedure . . . 135
6.2.10 Passive listening task . . . 135
6.2.11 Match-Mismatch task . . . 136
6.2.12 Behavioral procedure . . . 136
6.2.13 EEG data acquisition . . . 136
6.2.14 EEG data processing . . . 136

6.2.15 EEG data analysis . . . 137
6.2.16 Timeline . . . 138

7 General Discussion 147

7.1 Overview . . . 147
7.2 Main results . . . 148
7.3 Our studies and models of embodied semantics . . . 154
7.4 Conclusion . . . 158
7.5 Perspectives . . . 158

References 159

Chapter 1

The theoretical context

1.1 General Introduction

Our ability to learn and process language is perhaps the most remarkable and complex cognitive skill we possess. So complex that, despite decades of cognitive neuroscientific research, the nature of semantic representations is still a matter of debate. Using language requires, first, learning to associate concepts to linguistic labels and second, associating linguistic labels to concepts during comprehension. Language learning and comprehension are hence key aspects of human communication and, at the heart of these two inextricable processes, lies meaning. Cognitive psychology would have it that our linguistic repertoire is made up of abstract, amodal linguistic symbols, along with a set of rules that define how they can be combined (Fodor, 1980; Hummel, 2010). And yet the question remains, how exactly are semantic representations formed and retrieved? And how does the human brain associate the acoustic word [guitar] with the concept GUITAR, considering the completely arbitrary relationship between the two (Saussure, 1916; Shapiro, 2011)?

According to classic amodal theories, cognition is a computational process that creates meaning from perception and for action through the manipulation of mental symbols (Fodor 1987, 1994, 1998; Landauer & Dumais, 1997). This process is often described through the "sandwich model", according to which sensory-motor systems simply perceive input information from the world and produce output action on the world (Hurley, 1998). Sandwiched in between the two is cognition, which transforms perceived information into amodal symbols, links these symbols to relevant information in our semantic memory and performs operations on them for output. Knowledge is therefore stored in a semantic memory system isolated from sensory-motor processes (Barsalou, 1999) while a "cognition central" manipulates amodal symbols to create meaning from input and for output (Wilson & Foglia, 2017). Therefore the semantic network (Collins & Quillian, 1969) and distributional models of semantics (Landauer & Dumais, 1997), for instance, assume that concepts are represented by abstract and amodal symbols that are defined by other abstract, amodal symbols.

1.1.1 The symbol grounding problem and embodied semantics

A counter-argument to amodal theories is proposed by the symbol grounding problem, the question of how language acquires meaning without being grounded in something other than meaningless symbols (Glenberg, Havas, Becker & Rinck, 2005). This dilemma has been described using Harnad's (1990) Chinese-Chinese dictionary thought experiment, itself inspired by Searle's (1980) Chinese room argument. In this thought experiment, a woman must learn to speak Chinese using only a Chinese-Chinese dictionary. All symbols are defined by other symbols, none of which she understands, themselves defined by other unknown symbols, rendering it impossible for her to grasp the meaning of any of these symbols. Simply put, symbols must, at some point, relate to the real world and be grounded in sensory-motor experience (Hauk & Tschentscher, 2013). This is precisely the theoretical underpinning of embodied semantics, whose premises challenge many of the fundamental beliefs of cognitive research (Wilson & Foglia, 2017). Embodied semantics claims that linguistic forms are grounded in our body's system of perception and action planning (Barsalou, 2003; Barsalou, 1999). An often-cited justification for this is the "correlational learning principle", or the idea that the co-occurrence of action-perception and meaning results in the common firing of neurons that form "embodied referential semantic circuits" to support meaning representation (Pulvermüller, 2013). Neural connections are created between the two, forming distributed neural networks, which go on to subserve semantic processing (Pulvermüller, 1999). In short, "What fires together, wires together" (Hebb, 1949). Similarly, Barsalou's theory of perceptual symbols claims that during perceptual experience, the sensory-motor system captures modal states associated with experiences and encodes multimodal representations (Barsalou, 1999). As such, the acoustic form "guitar" gains meaning when tied with sensory experiences linked to the object, such as seeing, listening to and playing a guitar. When the lexical entry [guitar] is processed, the listener simulates the perceptual symbols, or experiences, associated with the concept GUITAR.

1.1.2 Semantic representations

The main distinction between amodal and embodied theories, and even within embodied theories, lies in how they regard the nature of internal mental representations, the meaning-containing substrates of cognitive processes (Aravena et al., 2014). These theories differ as to the format of internal representations, or what Machery (2016) refers to as their “representational code”. While embodied theories that acknowledge the existence of mental representations claim that motor-perceptual and linguistic mechanisms share, at least partially, a representational code, amodal theories regard the two as completely separate and fully independent from one another (Machery, 2016; Wilson & Foglia, 2017). In this view, modality-specific representations are transformed into amodal, mental symbols of their referents in the world. The symbol for the concept TABLE, for instance, is used for different types of tables and its relation to perceptual states involved in experiences with specific tables is completely arbitrary (Levelt, Roelofs & Meyer, 1999; Saussure, 1916). In sum, amodal theories claim that representations used for conceptual knowledge and language are independent from the body and its experiences.

Recently, behavioral and neuroscientific studies within the framework of embodied and grounded cognition have provided evidence that mental processes may employ the same neural substrates as those used in perception and action (Barsalou, 1999; Clark, 2013; Coello & Fischer, 2016). This has challenged the idea of amodal representations and suggested that internal representations may instead be modal. Radical embodied or non-representational theories such as the dynamic systems theory go as far as doing away with internal representations altogether (Beer, 1990, 2003; Chemero, 2013; Gibson, 1979; Thelen & Smith, 1994). Their general claim is that that there is no need for abstract symbolic concepts, as the direct influence that the body receives from its surroundings and their affordances is sufficient for cognition (Wilson & Foglia, 2017). That said, it would seem very challenging, for the time being, not to incorporate mental representations into theories of cognition when processes such as planning and imagining occur with no “help” from the immediate environment (Clark, 1997). Indeed, most embodied and grounded theories admit the existence of mental representations, only they regard these representations as being highly influenced by or even dependent on sensory-motor processes. Barsalou’s theory of perceptual symbols, for instance, is in agreement with classic models in that symbols are essential for cognition (Barsalou, 1999). However, he considers that perceptual and cognitive mechanisms share representations and that the symbols in question are therefore modal (Barsalou, 1999, 2008; Coello & Fischer, 2016; Kiefer & Pulvermüller, 2012). One of the pillars of this theory is that perceptual simulations can be involved in cognitive functions such as language processing. What this means is that representations have “activation patterns” that include perceptual state information originating in sensory modalities (Aravena et al., 2014; Barsalou, 2020; Wilson & Foglia, 2017). According to Barsalou (2009), what we know about a category is contained in a simulator. This simulator performs a simulation, or an internal enactment, during mental operations such as language processing for each specific occasion when meaning must be retrieved (Barsalou, 2020). The theory makes use of “association areas”, which integrate modality-specific representations (Barsalou, Dutriaux & Scheepers, 2018; Simmons & Barsalou, 2003). According to this theory, the features of a lexical item like [guitar] would be activated along with the relevant associated areas that integrate specific modalities involved in one's experiences involving guitars (Barsalou, 2020).

Proponents of a symbolic semantic system have suggested that a semantic "hub" acts as meeting point for the different semantic properties of a word (Pulvermüller, 2013; Garagnani & Pulvermüller, 2016) and, using evidence from neuroimaging and neuropsychological studies, they have proposed cortical regions for this multimodal hub (Binder & Desai, 2011; Geranmayeh, Leech & Wise, 2016; Patterson & Lambon Ralph, 2016; Price, Peelle, Bonner, Grossman & Hamilton, 2016; Simmons & Martin, 2009; see Pulvermüller, 2013 for a detailed description). However, as Pulvermüller (2010, 2013) points out, these studies all propose different regions as possible "hubs" for supporting semantic processes, in contrast with the idea of a general semantic hub. Furthermore, it has been shown that lesions affect the processing of distinct semantic categories differently, depending on their location (Damasio & Tranel, 1993; Gainotti, 2006; Warrington & McCarthy, 1983). This evidence deviates from the concept of an amodal semantic system that is limited to a specific cortical area and instead suggests that a much more distributed network underlies semantic representation (Pulvermüller, 2013). Specifically, Pulvermüller posits the existence of multimodal convergence zones for sensory, motor and affective information in the the following regions: the prefrontal, posterior parietal and anterior inferior cortex, as well as the posterior temporal cortex, among others. Furthermore, the observation that the different hubs that have been suggested to underlie semantic processing are category-specific (Pulvermüller, 2013) and overlap or are adjacent to sensory-motor areas, has lead to a mechanistic approach that links semantic and action-perception processes (Kiefer & Pulvermüller, 2012; Pulvermüller, 2005).

1.1.3 The complexity of putting embodied semantics to the test

At first glance, embodied semantics may appear to be a rather straightforward description of the semantic representation system; however, providing evidence to support embodied conceptual representations has proven complex. Most studies in this field have focused on showing a contribution of the motor and premotor cortex to action representations by 1) observing motor activation during action language processing or 2) inducing semantic and motor processes in a sequential or simultaneous manner to reveal either facilitation or interference. Nevertheless, attempting to show motor-semantic interactions through the overlap between the processes involved in comprehending action language and preparing or executing movement poses methodological challenges. One of these is attaining sufficient temporal precision to test the true nature of motor-semantic interactions. Behavioral studies have indeed revealed interference between semantic and perceptual-motor information (Glenberg &

Kaschak, 2002; Boulenger et al., 2006; Meteyard, Bahrami & Vigliocco, 2007). However, as behavioral measures occur late in language processing (i.e. after lexical access), both facilitation and interference effects may reflect different processes such as imagery (Mahon & Caramazza, 2008). Furthermore, the majority of neuroimaging evidence in favor of embodied semantics is provided by functional magnetic resonance imaging (fMRI) studies showing either sensory brain region activations for imageable words or somatotopical activation along the motor strip during action verb processing (Boulenger, Hauk & Pulvermüller, 2009; Esopenko et al., 2012; Hauk, Johnsrude & Pulvermüller, 2004; Raposo, Moss, Stamatakis & Tyler, 2009; for a review, see Aziz-Zadeh & Damasio, 2008). These activations are generally interpreted as evidence that sensory or action language automatically triggers the areas involved in sensorial processes or action execution. However, although precise when it comes to localization, fMRI is limited in temporal resolution and hence ambiguous as concerns the processing stage during which these activations occur (Hauk & Tschentscher, 2013; Mollo, Pulvermüller & Hauk, 2016).

Semantic processes have been shown to occur very rapidly; thus there is a significant risk of confounding a behavioral or cortical response associated with semantic processes with any number of processes that may occur in parallel with or after the relevant time frame (Hauk, 2016). Electroencephalography (EEG) offers much higher temporal resolution and can therefore better pinpoint cortical activity that may contribute earlier to semantic processes, beginning around 150-200 msec after word onset (Amsel, Urbach & Kutas, 2013; Barber & Kutas, 2007; Moseley et al., 2013, Pulvermüller, 2005). Importantly, an EEG-magnetoencephalography (MEG) study revealed motor-semantic interactions, as indexed by effector-type and action word category congruency effects, as early as 150 msec post-stimulus (Mollo et al., 2016). That said, later activation can also be associated with semantic processes that precede mental imagery. Hauk & Tschentscher (2013) argue that, although there is no consensus yet as to the exact time course of lexico-semantic processes, it is probably safe to assume that activity occurring before 400 msec is not related to mental imagery. EEG has thus proven well-suited for measuring motor-semantic interactions that occur early enough in language processing that they could be interpreted as showing that motor processes influence semantic processes and/or vice versa. Importantly, EEG is used to study phase-locked event-related potentials (ERPs), but also cortical activity that is not phase-locked, such as variations in neural oscillations in different frequency bands. These variations, referred to as event-related synchronization (ERS) and desynchronization (ERD) can be examined through time-frequency analysis (Pfurtscheller & Lopes da Silva, 1999). Throughout the work presented in this dissertation, we have capitalized on both ERPs and time-frequency analysis as a means to acquire cortical measures of motor-semantic interactions.

1.1.4 Contextual language processing and embodiment

A cognitive process — be it emotional, language-related, motor-related, or sensory-related — always takes place in the context of other processes. It is hence influenced by them and, in turn, has an influence on them (García & Ibáñez, 2016). García and Ibáñez refer to the holistic nature of neurocognition as "sensus communis", a dynamic convergence of different processes that ensue from external and internal events and result in experience. As such, when studying one cognitive process, other co-occurring processes must be taken into account for the process of interest to be fully understood. This, however, is easier said than done as it entails accounting for a series of very complex mechanisms. Within the framework of embodied semantics, the studies presented in this dissertation were mainly concerned with accounting for sensory-motor and language processing mechanisms in the presence of one another and a considerable effort was made to enrich contextual cues through linguistic materials, as well physical and environmental factors. When investigating language processing in the context of motor processes, and vice versa, ecological validity is of great importance (Peeters, 2019). It is also true that cognitive studies, especially those that involve cortical measures using techniques such as EEG, require substantial experimental control, both when it comes to linguistic elements and physical/environmental aspects. Therefore, in the embodied semantics line of research, a delicate balance must be attained between ecological validity and experimental control.

This presents a considerable challenge. When using cortical measures, it is impossible, for the moment, to provide rich contextual cues in all these domains at once, while maintaining experimental control. This should, however, be an ultimate goal for language processing research and furthering this possibility was one of the main concerns of this dissertation. The reason behind this is simple, but often overlooked: natural language learning and processing is *interactive* and *contextual* (Atkinson, 2002; Peeters, 2019). Associating an auditory word to its referent, whether for the first or for the umpteenth time, is a multimodal experience that generally takes place in sensorially rich contexts. Moreover, in real life, we use language while interacting with our environment through both observing and performing action. Our bodies and the environment surrounding us are hence intrinsically linked to how we process language. What remains unknown as regards learning and processing linguistic semantic content can best, and possibly only, be revealed through paradigms that truly take this into consideration.

The studies presented in this dissertation follow a progression of increasing contextual complexity and ecological validity. Study one (chapter 2) provided linguistic context by having participants learn second-language (L2) words in the context of sentences. We also designed interactive computer-games to begin exploring the idea of contextual interaction during language learning. In study two (chapter 3), we used a traditional ACE paradigm, and, once again, presented action language in the context of sentences, this time in the participants' native language. We designed an answer box that required real action; participants fully stretched their arm forward to answer "away" from themselves and brought their arm in to answer "toward" themselves. This was a first step in inducing contextual movement during action language processing. In study three (chapter 4) we took a leap toward ecological validity with our use of an immersive virtual reality (VR) CAVE (Cave automatic virtual environment) and a finger-tracker to provide a context-rich environment in which participants heard words and manipulated 3D objects using 12 different, corresponding actions. Finally, in the registered report presented in chapter 6, we developed a virtual reality paradigm to investigate L2 learning using an Oculus headset and hand controllers. In this projected study, participants will once again perform real actions on virtual objects in an immersive 3D virtual reality environment, this time while learning L2 action verbs. Very seldom have controlled experimental studies investigating language processing and language learning provided such rich linguistic and physical environments.

1.2 Electroencephalography to measure language processing

1.2.1 Event-related potentials (ERPs)

Neurolinguistic studies have as their main goal the comprehension of the cognitive correlates of both language representation and processing through the identification of their neural substrates (Pulvermüller, 2010). Event-related potentials (ERPs) reflect brain activity resulting from a multitude of post-synaptic potentials generated by the brain in response to sensory, cognitive and motor events (Luck, 2014). They are extracted by averaging the electroencephalography (EEG) signal obtained across trials and then across participants, thus filtering out background activity and retaining robust brain activity that results from processing the stimulus (Morgan-Short & Tanner, 2013). ERPs have been widely used in language studies to measure cortical activity time-locked to visual and auditory stimuli, which have contributed to our understanding of the neural mechanisms underlying language processing. ERP components are characterized by their onset latency, amplitude and/or topography. Differences in any or all of these measures as a function of experimental conditions have been used to provide support for various theories of linguistic representation and processing (Morgan-Short & Tanner, 2013).

Particularly pertinent in language processing studies is the N400 component, a negative-going wave usually observed between 300-500 msec after stimulus onset, at central and posterior electrodes over both hemispheres. The N400 has been shown to reflect semantic integration such that a greater N400 amplitude is attributed to processing difficulty that results from attempting to integrate a new element within an existing semantic context (Holcomb, 1993; Kutas & Federmeier, 2011; Kutas & Hilliyard, 1980). As a result, during language processing, the more easily a word is integrated into its semantic context, the greater the attenuation of the N400 amplitude (Brown & Hagoort, 2000; Collins & Loftus, 1975; Hahne & Friederici, 2002; Kutas & Hilliyard, 1980). During L2 processing, N400 peak amplitude is often delayed (Ardal, Meuter, Muldrew & Luce, 1990; Hahne, 2001; Ojima, Nakata & Kakigi, 2005; Newman, Tremblay, Nichols, Neville & Ullman, 2012; Weber-Fox & Neville, 1996; but see Bowden, Steinhauer, Sanz & Ullman, 2013). Also, in terms of topography, instead of central-parietal sites, newly learned words can yield N400 effects over frontal electrode sites (Elgort, Perfetti, Rickles & Stafura, 2015). The N400 component likewise varies in match-mismatch tasks that present written or oral lexical items paired with matched or mismatched images (Cummings, Čeponiené, Dick, Saygin & Townsend, 2008; Hamm, Johnson & Kirk, 2002). Mismatched pairs generally result in greater N400 amplitude compared to matched pairs (Hamm et al., 2002). We measured semantic processing via the N400 in our first study (chapter 2), in which we examined L2 vocabulary learning; in our second study (chapter 3) in sentential contexts using an ACE paradigm; and finally in our projected virtual reality (VR) L2 learning study (chapter 6). We used a match-mismatch task in both our first study (chapter 2) and in our registered report (chapter 6) to examine vocabulary learning. In both of these studies, we used a pre/post-test design to measure differences in the N400 amplitude as a function of the type of pair (match vs mismatch) of auditory words preceded by either static or dynamic images.

1.2.2 Time-Frequency ERSP to measure motor activation during language processing

Another way of examining the EEG signal is by observing changes in neural oscillations in different frequency bands through time-frequency analysis. Recent years have seen a significant increase in time-frequency studies on different aspects of language processing (Ghitza, 2017; Giraud & Poeppel, 2012; Kösem & van Wassenhove, 2017; Rommers, Dickson, Norton, Wlotko & Federmeier, 2017; Litcofsky & Van Hell, 2017; for a review see Hauk, Giraud & Clarke, 2017). Despite the temporal precision afforded by ERPs, they are limited when it comes to measuring motor processes, as they use phase-locked linear averaging methods, which do not take into account the great deal of motor cortex signal that is not phase-locked to the stimulus (Pfurtscheller & Lopes da Silva, 1999; Vukovic & Shtyrov, 2014). Post-stimulus spectral estimation methods have been used to measure both phase-locked and non-phase-locked activity in the motor cortex, in the time interval following stimulus onset compared to a pre-stimulus baseline (for a review see Hobson and Bishop, 2016). Event-related spectral perturbation (ERSP) groups together event-related desynchronization (ERD) and synchronization (ERS), power suppression and enhancement respectively, and calculates the power spectrum on a trial-by-trial basis (Grandchamp & Delorme, 2011; Makeig, 1993; Makeig, Debener, Onton & Delorme, 2004). Much of the research on motor activation using ERSP has focused on mu (8-13Hz) and beta (13-30Hz) rhythms, as their suppression reflects synchronized patterns of electrical activity recorded over the sensory-motor cortex, associated with performing and observing movement (Caetano, Jousmäki & Hari, 2007; Koelewijn, van Schie, Bekkering, Oostenveld & Jensen, 2008; Niccolai et al., 2014; Pfurtscheller & Lopes da Silva, 1999; Pineda, 2005; Wamain, Sahaï, Decroix, Coello & Kalénine, 2018). Motor imagery has likewise been linked to a crease in the mu rhythm compared to a baseline interval (Matsumoto et al., 2010).

Recently, a growing number of studies have used time-frequency analysis to measure motor activation during language processing. Reading single verbs related to the body caused soma topical mu ERD (Niccolai et al., 2014). Furthermore, processing action-related sentences caused mu and beta ERD; this was interpreted as showing motor resonance during the retrieval of lexical-semantic information as opposed to a post-lexical imagery (van Elk, van Schie, Zwaan & Bekkering, 2010). Generally speaking, greater mu rhythm suppression has been observed for action language compared to abstract language (Alemanno et al., 2012; Moreno et al., 2015). The studies presented in chapters 3 and 4 use time-frequency to directly measure motor resonance during language processing by observing mu and beta desynchronization time-locked to the critical verb. As described above, the study in chapter 3 focuses on action sentence processing using an ACE paradigm to investigate how action and meaning congruency influence one another. The study presented in chapter 4 examines motor resonance during single verb processing in an immersive and interactive virtual reality environment. Our registered report in chapter 6 will also employ time-frequency to compare motor activation during verb processing pre and post-training.

1.3 Using the ACE to measure interactions between motor and linguistic processes

The Action-sentence compatibility effect (ACE) paradigm, originally developed by Glenberg and Kaschak (2002), provides a means to measure the impact of language processing on motor processes. Participants make a sensibility judgment after hearing or reading a sentence that describes an action while (or before or after) performing a physical action that is either compatible or incompatible with the action described. Glenberg and Kaschak (2002) found that arm movements were facilitated when they were compatible with the direction described in auditory sentences. For example, when participants responded to sentences by moving their arm away from their body, their response times were faster for compatible sentences ("Close the drawer"). The authors attributed this ACE to participants performing a mental simulation of the sentence in order to comprehend it.

Since then, variations of the ACE paradigm have shown a clear ACE using different hand and arm movements (Diefenbach, Rieger, Massen & Prinz, 2013; Dudschig, de la Vega & Kaup, 2014; Zwaan & Taylor, 2006). However, more recent studies have revealed that factors such as the task (Sato, Mengarelli, Riggio, Gallese & Buccino, 2008), linguistic context (Aravena et al., 2012) and linguistic perspective (Gianelli, Farnè, Salemme, Jeannerod & Roy, 2011) have a strong impact on whether or not an ACE is found. Furthermore, the direction of the ACE has been shown to change contingent on response timing (Boulenger et al., 2006) and movement type (Shebani & Pulvermüller, 2018). When considering counter-arguments to embodied cognition, especially as regards the non-functionality of motor activation during language processing, timing is of utmost importance. Otherwise stated, whether motor activation is contingent on meaning activation or subsequent to it is crucial to the argument that motor processes play a role in linguistic processing.

1.3.1 The neural correlates of the ACE

A handful of ACE studies have used cortical measures to investigate the question of timing as concerns the overlap of motor and linguistic processing (for a review see García & Ibáñez, 2016). EEG is an ideal tool for tackling this issue. Indeed, combining the ACE paradigm and cortical measures (EEG) is an efficient means for measuring the overlap between motor and semantic processes at the cortical level with millisecond precise timing, and hence for answering many questions regarding embodied semantics. Using an ACE paradigm and ERPs, Aravena and colleagues manipulated action and language compatibility to examine motor-language interactions (Aravena et al., 2010). The rationale behind this study was that if semantic representations are mapped onto sensory-motor systems, it follows that language comprehension should engage sensory-motor processes. Specifically, if language understanding and sensory-motor processes share neural resources, then their cooperation would result in language facilitating compatible movement and, vice versa, with incompatible movement interfering in language processing. The authors were therefore interested in examining the brain markers of what they called "the bidirectional impact" between linguistic and motor processes. Theirs was the first study to combine an ACE paradigm and ERP measures with this goal. They found a greater deflection of an N400-like component time-locked to the critical verb for incompatible compared to compatible trials and greater positivity in motor potentials (MP, -90 to 50 msec and RAP, 200 to 300 msec) for compatible versus incompatible trials. The authors interpreted these bidirectional motor-semantic effects as evidence for a strong ACE.

Aravena et al.'s study (2010) goes beyond behavioral ACE findings as regards supporting embodied semantics in that it reveals an early semantic priming facilitation on motor performance. The authors argue that the timing of this effect indicates an ongoing motor-language interaction as opposed to a post-lexical effect. Indeed, although motor resonance has been shown to affect language processing through motor-semantic compatibility facilitation or interference in behavioral ACE studies, examining the temporal dynamics of the ACE with greater precision makes for a stronger argument for semantic-to-motor facilitation. Importantly, although previous behavioral ACE studies (Glenberg & Kaschak, 2002; Diefenbach et al., 2013; Dudschig et al., 2014; Zwaan & Taylor, 2006) indicated that semantic processes primed motor responses, they did not provide evidence for the opposite. Aravena and colleagues' study, on the other hand, was concerned with a bidirectional effect and is one of the rare electrophysiological studies to also investigate the effect of motor processes on language comprehension. The bidirectional semantic-to-motor and motor-to-semantic effects in this study do not prove that motor activation has a causal role in language processing. They do, however, evidence a dynamic and robust interaction between the two processes that goes beyond what behavioral ACE studies had shown previously. In this sense, the results of this work contradict an abstract and amodal view of language. If the linguistic symbols we manipulate for cognition were truly free of modal information, compatibility between language and performed action would not have such an impact on both motor performance and semantic processing.

1.3.2 Further exploring the neural correlates of the ACE using EEG

As has been noted time and time again, we are at a point in the embodiment debate where rather than pursuing the embodied versus disembodied argument, it has become much more productive to ask *how* and *when* language processes are embodied (Chatterjee, 2010; Mollo et al., 2016). For example, as regards *how*, whereas language-to-motor effects have been widely shown, there is significantly less evidence of motor-to-language effects. Within Barsalou's simulation framework (1999) as well as according to Pulvermüller's associative learning hypothesis (2005), it would stand to reason that the effect would go both ways. Also, considering that the vast majority of embodied semantics studies do not use high time course precision methods (EEG, MEG), the *when* question still remains mostly unanswered. Examining early effects, both motor and semantic, as did Aravena et al. (2010), is crucial for further refining the relationship between motor and semantic processes.

In the ACE study we present here (chapter 3) our initial aim was to replicate the effects shown in Aravena et al.'s study, but using materials and a task that resemble more closely Glenberg and Kaschak's original study (2002) whereby participants were asked to judge the acceptability of action sentences by performing either a compatible or an incompatible action. However, whereas Glenberg and Kaschak used imperative second-person sentences ("Close the drawer"), like Aravena et al. we used third-person sentences (*Emilie a pris son verre de vin et l'a bu* [Emilie picked up her glass of wine and drank it]) to avoid imagery. Crucially, Aravena et al.'s study required participants to perform a repetitive action (a button press using a closed or open hand) in a situation where they could prepare for this action early in sentence processing; in our version of the paradigm, participants could only know in which direction they would answer at the very end of the sentence. As in Aravena et al.'s study, we examined both motor preparation and semantic processing using ERPs. In an effort to measure early motor activation during action language processing and present a more complete picture of motor-semantic interactions, our study also used time-frequency measures. Specifically, we examined mu and beta desynchronization compared to baseline, relative to verb onset. To our knowledge, ours is the first study to combine these two measures to examine the ACE. call Siri

1.4 Adult L2 learning

Adult word learning in a second language (L2) differs from child first-language acquisition in that children learn their first language while acquiring general knowledge about their environment, whereas adults already possess much of this knowledge (MacWhinney, 2009; Rodríguez-Fornells, Cunillera, Mestres-Missé & Diego-Balaguer, 2009). Adult word learning processes are more varied and success is contingent on a myriad of factors such as the learner's linguistic background, motivation, cognitive ability and other individual differences (Carroll, 1993; Rodríguez-Fornells et al., 2009). Learning modality is also an important element; exposure and practice can include rich multimodal stimuli, images, videos, auditory information and written text. Finally, context, which can range from highly structured classroom instruction to full immersion, can impact learning outcomes (Cook, Long & McDonough, 2008). In addition, whereas most studies that investigate the cortical processes underling lexical acquisition have focused on infants and children (Friedrich & Friederici, 2008; Friedrich, Wilhelm, Born & Friederici, 2015; Childers & Tomasello, 2002), the majority of the words we know, whether in an L1 or in an L2, are learned after early childhood (Borovsky, Elman & Kutas, 2012). Importantly, the way in which we encode words is thought to change between childhood and adulthood (Borovsky et al., 2012). Given the complexity and variability of adult word encoding, much remains to be explored as concerns the cognitive processes behind it.

Studies that capitalize on the learning modality and context, for example, can greatly further our understanding of what is at play during adult word learning. In this framework, numerous behavioral studies have examined the effect of immersion for L2 learning, through game-playing.

1.4.1 Games for interactive L2 learning

Linguistic immersion has been associated with significant gains in L2 word learning compared to traditional classroom learning (Freed, Segalowitz & Dewey, 2004). The vast majority of language teachers support cultural immersion, with an emphasis on interaction as a means to succeed in acquiring a new language (Young et al., 2012). An interesting experimental approach to increase immersion and interaction during language learning is digital game-playing. On and offline computer-games range from simple text games to virtual games that can involve several players concurrently (Sylvén & Sundqvist, 2012). According to Gee (2012), compared to classroom learning, playing video games offers an experience that is closer to *in situ* learning as it “situates the senses” by associating words not to other words but to images and actions. He notes that, much like human dialogue, this gives participants a deeper understanding and practice as concerns the L2 (Gee, 2012).

In the last twenty years, several studies have measured the benefit of playing computer-games for L2 learning. Generally speaking, they have found correlations between digital game-playing and improvements in L2 vocabulary acquisition, writing, classroom participation and verbal fluency (Chik, 2014; Kuppens, 2010; Peterson, 2010; Ranalli, 2008; Sundqvist, 2009; Sylvén & Sundqvist, 2012; Young et al., 2012). These improvements have been linked to the immersive environment provided by digital games, especially Massively Multiplayer Online Role-playing games (MMORPG) or 3D games, such as *World of Warcraft* (WoW), *Second Life* or *Active Worlds*.

In the study described in chapter 2, we created flat-screen computer-games to teach adults a small L2 vocabulary. We opted for a computer-games teaching strategy because, as described above, this has been shown to be more engaging and helpful for L2 learning than traditional teaching methods. We also considered this a first step towards more immersive and embodied learning methods, which we developed in chapter 6. Four games designed to teach L2 vocabulary to absolute beginners were created in collaboration with the Mediterranean Virtual Reality Center (CRVM). In order to ensure their effectiveness, the following linguistic notions were taken into consideration: 1) Implicit (Exposition game) and explicit (Segmentation game) learning are both important in vocabulary acquisition 2) L2 learners often find it difficult to segment the auditory sequence into meaningful lexical units (Segmentation game) 3) Both L1 and L2 learners learn through opposition (Exposition and Memory games) 4) L2 learners often have difficulty treating new vocabulary in sentential context as compared to isolated words (Memory game). Visually, our games were very simplistic as they displayed a stick figure performing actions on objects depicted by black and white line drawings. However, as further described in chapter 2, they proved very efficient. In addition to computer-based learning, virtual reality, as will be developed in our final study (chapter 6) can provide a unique learning environment.

1.4.2 Embodied semantics and language learning: behavioral evidence

Embodied semantics offers a valuable theoretical framework for examining adult word learning, especially as regards the role of the body in encoding new information (Coello & Bartolo, 2013). According to the Hebbian theory of associative learning, when lexical items are acquired along with action, neurons in cortical regions involved in language processing and those involved in action planning and execution are co-activated and form neuronal assemblies (Hebb, 1949; Pulvermüller 1999). It has been suggested that this synchronous activation during encoding is the reason behind motor and language processing interactions (Pulvermüller, 1999, 2005; Tomasello, Garagnani, Wennekers & Pulvermüller, 2018). Along similar lines, Zwaan and Madden's (2005) theory of experiential traces claims that when we interact with the environment, this forms 'experiential traces' in our brains, which become associated with co-occurring linguistic labels. Hence when we later come in contact with the same linguistic labels, our experiential traces are automatically reactivated. In line with these theories, in the last forty years behavioral studies have amply documented that physical movement that is congruent with new linguistic content supports retention, a phenomenon often referred to as the "enactment effect" (Engelkamp, 1980; Engelkamp & Krumnacker, 1980; Engelkamp & Zimmer, 1984). For example, Engelkamp and Krumnacker (1980) showed that verb phrases such as "shuffle the cards" were better memorized when learners performed representative gestures during learning compared to either watching someone else perform the action, imagining the action or simply listening to the sentence. More recently, encoding with action has been termed "embodied learning". The general consensus is that truly embodied learning involves "self-performed" or "self-generated" action that is congruent with learned content (James & Bose, 2011; James & Swain, 2011; Johnson-Glenberg, 2017, 2018; Johnson-Glenberg & Megowan-Romanowicz, 2017). In fact, recent embodied learning taxonomies propose different levels of embodiment for education based on how physically engaged learners are, as well as the congruency between gestures and the content being learned (Johnson-Glenberg & Megowan-Romanowicz, 2017; Skulmowski & Rey, 2018). For instance, physical activity facilitated the integration of sung melodies (Wakefield & James, 2011) as well as mathematical (Kontra, Lyons, Fischer & Beilock, 2015) and scientific principles (Johnson-Glenberg & Megowan-Romanowicz, 2017; Johnson-Glenberg, Megowan-Romanowicz, Birchfield & Savio-

Ramos, 2016) more than verbalization alone. Moreover, children who gestured while learning a new mathematical concept showed better long-term retention compared to those who only verbalized the new information (Broaders, Cook, Mitchell & Goldin-Meadow, 2007; Cook, Mitchell & Goldin-Meadow, 2008).

Embodied learning also supports language learning (Macedonia & Knösche, 2011; Mayer, Yildiz, Macedonia & von Kriegstein, 2015; de Nooijer, van Gog, Paas & Zwaan, 2013; Tellier, 2008). For example, simply pointing at or touching objects during L2 word learning led to the association between sensory-motor experiences (the location of an object) and the novel word (Öttl, Dudschig & Kaup, 2017). Furthermore, both children and adults were more successful at learning L2 words when they used representative or iconic gestures (McNeill, 1992) compared to control conditions using images (Tellier, 2008; Macedonia & Knösche, 2011). Words describing object manipulation (but not abstract words or words describing locomotion) were better encoded when participants imitated a physical movement during encoding and retrieval compared to a control condition (de Nooijer et al., 2013). Finally, long-term retention is improved for words learned with action compared to control conditions (Macedonia & Klimesch, 2014; Mayer et al., 2015). The great amount of behavioral studies evidencing that retention is improved through embodied learning supports the Hebbian and experiential trace theories described above (Hebb, 1949; Pulvermüller 1999; Zwaan & Madden, 2005) and suggests that encoding lexical items in association with action likely links semantic meaning to motor traces.

1.4.3 Learning studies to explore the neural correlates of embodied encoding

One way of exploring the relationship between sensory-motor experience and language representation such as described by the associative learning model (Pulvermüller, 1999) or the theory of experiential traces (Zwaan & Madden, 2005) is to teach learners novel words along with physical action and then test whether these words produce motor activation using time-frequency analysis. Fargier and colleagues carried out a learning paradigm during which pseudo-words were learned in two conditions: in association with self-performed actions and with abstract animations (2012). After two training sessions, accuracy was at ceiling for both conditions. Post-training centrally distributed mu desynchronization revealed greater motor activation for words learned in the motor action condition compared to the control condition on day one of learning. However, on day two of training a fronto-central distribution was observed for mu desynchronization, which the authors interpreted as not directly reflecting motor activation but activity in a “convergence zone” between motor and language structures (Fargier et al., 2012).

Bechtold and colleagues had participants learn novel names for novel tools either through visual or manual exploration and found greater mu and beta ERD post-training for words learned through haptic manipulation (Bechtold, Ghio, Lange & Bellebaum, 2018). However, they also found comparable motor activation for non-tool related familiarized pseudo words, leading them to conclude that differences in mu and beta ERD during the processing of these words were possibly due to the suppression of motor activation for processing words that only have visual features. The above results do not provide a clear-cut answer to whether words learned with action directly reactivate sensory-motor information. What's more, the question remains whether there is a correlation between motor activation after embodied word learning and improved learning outcomes. We will further examine this question in our final projected study (chapter 6), in which participants will learn an L2 vocabulary of action verbs in an interactive virtual reality environment. During training, participants will manipulate a virtual object and perform specific actions (a different action for every action verb) compared to a control condition in which participants will simply point to the virtual object during training. The study uses EEG to measure whether embodied learning leads to greater motor activation (time-frequency measures) and improved retention (ERPs) post-training compared to the control condition.

1.5 Virtual Reality

1.5.1 VR: a naturalistic environment in which to measure language processing

One of the major challenges faced by embodied semantics studies is providing ecological validity while maintaining experimental control. Embodied theories place great importance on environmental and physical contexts, thus calling for multimodal experimental protocols that are closer to real life than those generally used in psycholinguistic studies (Tromp, Peeters, Meyer & Hagoort, 2018; Peeters, 2019). Importantly, embodied semantics is strongly linked to theories of grounded or situated cognition, which claim that the environment plays a crucial role in cognition. Accordingly, language processing and language learning environments, as well as interactions with such environments, are an integral part of cognition (Atkinson, 2010; Black, Segal, Vitale & Fadjo, 2012). On the other hand, given the need for control, experimental protocols that examine "embodied" L2 language processing very seldom provide ecologically valid environments and fail to allow for naturalistic physical action. Movements are generally constricted and repetitive, and training occurs in isolated 2-D environments stripped of contextual everyday cues, making it difficult to apply the results obtained in these experiments to real-life language processing (Peeters, 2019). When one considers that real-world language processing generally occurs in visually rich environments and that physical cues have a strong influence on how language is understood, it becomes clear that the decontextualized 2-D conditions generally used in language processing studies limit their ecological validity (Knoeferle, 2015; Tromp et al., 2018). Studies that analyze the neural correlates of language processing and learning using techniques such as fMRI and EEG are especially limited when it comes to physical interaction and naturalistic movements, given their need to control for artifacts related to muscular activity (Luck, 2014).

Recently, language research has become increasingly invested in providing participants with more ecologically valid multimodal environments where they can engage in semi-natural movement, making for results that can generalize to everyday situations (Peeters, 2019). A multimodal approach to examining the neural underpinnings of language processing should allow for situated cognition in realistic environments, while giving experimentators the same amount of control as they would have in traditional experiments (Peeters, 2019; Tromp et al., 2018). In fact, the more real-world and situated a language processing environment is, the more physically implicated, and natural, participants will feel, and the more applicable the ensuing results will be to real-life processing (Peeters, 2019). Fortunately, modern technology makes it possible to study cognitive processes in richer contexts (Ladouce, Donaldson, Dudchenko, & Ietswaart, 2016). Virtual reality (VR), in particular, offers an environment in which to observe language processing and learning that is both controlled and ecologically valid (Peeters, 2019; Repetto, Cipresso & Riva, 2015; Tromp et al., 2018).

Peeters (2019) argues that VR eliminates the spatial divide between stimulus and participant and that it is an especially interesting approach for topics of a multimodal nature, and hence for investigating embodied semantics. There exist three basic types of VR environments: the flat-screen computer monitor, the head-mounted display (HMD) and the Cave automatic virtual environment (CAVE) system. The least immersive of these is the flat-screen computer monitor as it only occupies a small percentage of the participant's visual field (Repetto, 2014). HMDs provide an immersive experience as they visually isolate the participant from the real world, creating a sense of full immersion. A step further, the CAVE system allows participants to see their own bodies in motion, adding to their sense of presence (Moore, Wiederhold, Wiederhold & Riva, 2002), while surrounding them with 3 or 4 screens that provide a sensory illusion of a real environment (for a review see Bohil, Alicea & Biocca, 2011). Both the HMD and the CAVE technologies provide a sense of agency due to the participants' ability to move their arms and hands freely, as well as manipulate objects (Johnson-Glenberg, 2018). HMD and CAVE virtual reality technologies situate participants in 3-D environments, enhanced with visual and auditory stimuli. These environments can provide both a sense of immersion, or "a sensory-motor coupling between a participant and a virtual environment" and presence, "a psychological, attentional and cognitive state in which the participant, immersed within a virtual environment, behaves in accordance with the affordances provided by this environment […]" (Mestre, 2015, p.1). In ecologically valid VR environments, participants are free to interact with objects, while receiving real-time feedback via a graphic rendering system. Input tools (finger trackers, gloves, a mouse or joystick) are often used to record participants' movements (Burdea & Coiffet, 2003). Given the interactive and immersive nature of VR environments, it has been suggested that the sensory-motor system becomes more fully engaged than in traditional experiments, and that elicited responses are closer to what probably occurs in real life (Bohil et al., 2011).

1.5.2 Combining VR and EEG in language processing and learning studies

Within the framework of embodied semantics, the level of immersion, presence and ecological validity provided by VR makes it a very attractive methodology to pair with EEG for studying the interaction of motor and linguistic processing. In VR, multimodal sensory stimulation is fully controlled, rendering possible the direct observation of brain activity as a correlate of specific sensory input, whether visual or auditory, in an environment that allows for naturalistic actions. According to Peeters (2019, p.6), what is most promising about virtual reality as an experimental tool is that it will "shift theoretical focus towards the interplay between different modalities in dynamic and communicative real-world environments, moving beyond and complementing studies that focus on one modality in isolation".

Virtual reality has been used to replicate results from previous traditional language processing studies. For instance, predictive language processing was shown through language-mediated anticipatory eye movements in a VR environment (Eichert, Peeters & Hagoort, 2018). Another study compared human-human to human-avatar communication using a syntactic task and found similar priming effects (Heyselaar, Hagoort & Segaert, 2017). Relevant to embodied semantics, disrupting the visual feedback of participants' pointing trajectory in VR led to a delay in speech production, showing an interplay between hand movement and speech production mechanisms (Chu & Hagoort, 2014).

For now, however, very few studies have combined cortical measures and virtual reality to investigate language processing. A recent study used time-frequency analysis to investigate competition between gesture representations as participants perceived 3D objects in a VR environment. Results evidenced suppressed mu desynchronization when participants processed conflictual manipulatable objects in peripersonal space, possibly reflecting action selection processes (Wamain et al., 2018). Another study had unbalanced bilinguals name pictures for two virtual agents, using their L1 for the first agent and their L2 for the second agent (Peeters & Dijkstra, 2018). Behavioral measures revealed symmetrical switch costs across languages and greater mixing costs for the L1 compared to the L2. ERP results confirmed these results via a greater late central-parietal positivity for switch costs across languages and a late positive effect for the L1 and not the L2 when comparing non-switch to blocked naming trials. In an exploratory study, participants wore a head-mounted VR display as they listened to a sentence (“I just ordered this salmon”) and saw a virtual object that either matched (salmon) or mismatched (pasta) the object in the sentence (Tromp et al., 2018). A match-mismatch N400 effect was found for incorrectly orally labeled items compared to correctly labeled items. In sum, the combination of cortical measures using EEG and VR has clear potential for addressing questions related to language processing within an embodied framework.

In the work presented in this dissertation, we were especially interested in combining EEG with CAVE and HMD virtual environments to measure the effect of stimulating movement during linguistic processing, in order to provide novel and compelling insight into embodied semantics. This dissertation includes one published study and one registered report that make use of this seldom-used combination to investigate language processing and encoding. The study presented in chapter 4 combined a CAVE and EEG to measure motor resonance during single verb processing. In the registered report in chapter 6, participants learn a vocabulary in an L2 using an HMD and a handset to perform real actions on a virtual object in an immersive 3D environment. Motor activation and learning were measured using EEG, pre and post-training.

1.6 Completed Studies

1.6.1 Cross-linguistic gender congruency effects during lexical access in novice L2 learners: evidence from EEG

In chapter 2, we present a currently submitted study that explores the mechanisms underlying adult L2 word learning. We created digital games in collaboration with the Mediterranean Virtual Reality Center (CRVM) to teach a new lexicon to completely novice learners over four consecutive days. We observed behavioral and cortical changes during the very first phases of L2 learning. Thus, we were able to follow progression from zero knowledge to the understanding of simple declarative sentences involving an action verb and two nouns. Very few studies have examined the cortical response to the integration of novel words during the very early stages of learning (O’Neil, Lagarrigue, Newman & Frenck-Mestre, unpublished). We examined this question via a comparison of behavioral and cortical measures (N400) during an audiovisual match-mismatch performed both pre and post-training.

A crucial aspect of second-language (L2) acquisition is how a learner’s native language

(L1) affects the integration an L2 lexicon (Grosjean & Byers-Heinlein, 2018). According to the dominant models of bilingual lexical representation, such as the Bilingual Interactive Activation + model (Dijkstra & Van Heuven, 2002), lexical access is non-selective across a bilingual's L1 and L2. Grammatical gender, an intrinsic syntactic feature of lexical items in a number of languages, provides an interesting tool for measuring the influence of L1/L2 feature congruity on L2 learning. The first aim of our study was to examine how L1/L2 gender congruency might influence L2 lexical integration from the very initial stages of learning. Based on previous research showing cross-language gender congruency effects (Sá-Leite, Fraga & Comesaña, 2019) we hypothesized that if gender is shared across lexical entries, convergence of grammatical gender across the L1 and L2 would have immediate effects on learning. No specific training of gender was provided, hence any gender information acquired was exposure-based and implicit. All language exposure was oral and no L1 support was provided. An effect of gender congruency would suggest that L1/L2 differences interfere in L2 lexical integration during the early stages of L2 learning. Alternatively, if no effects of grammatical gender congruency are found, we could conclude that learners do not access L1 grammatical features such as gender online while treating L2 words, at least during the early stages of learning.

1.6.2 What role does motor activation play in action language processing? An EEG study

In our subsequent experiments, we took the interactive aspect of language processing a step further by exploring the role of the body in first language processing, and later investigated its influence in L2 learning. In these experiments, we continued to use behavioral measures and ERPs and added time-frequency analysis to measure motor activation during language processing.

In chapter 3, we report our second experiment, which examined the implication of physical movement for language processing. Using an Action-sentence compatibility effect (ACE) paradigm, we manipulated language and action congruency to measure the effect of such on language processing. Given that the direction of the ACE effect – whether facilitatory or inhibitory – is highly dependent on when motor preparation takes place in relation to action language processing (Aravena et al., 2010), we were interested in whether simultaneously planning a movement that was congruent with the action language being understood would interfere with semantic processing. Only one study has looked at the neural response to action language and movement congruency during overlapping motor and semantic processes (Aravena et al., 2010) and they found an enhanced N400-like amplitude for incompatible versus compatible pairs. They also observed greater motor preparation for congruent trials compared to incongruent trials, as indexed by ERPs associated with motor preparation (MP and RAP). In our study we likewise observed changes in the N400 time window to examine semantic processing. However, ERPs only show phase-locked activity and are hence not the ideal means of measuring activation in the motor cortex. Therefore, in order to obtain a more comprehensive view of possible interactions between motor and linguistic processing, we also used time-frequency analysis, which takes into account both phase-locked and non-phase-locked activity (Pfurtscheller & Lopes da Silva, 1999).

1.6.3 Motor Resonance during linguistic processing as shown by EEG in a naturalistic VR environment

To further explore embodied effects in language processing, we developed a novel paradigm combining EEG and a CAVE. In the published study presented in chapter 4, we took into consideration theories of situated and grounded cognition, which claim that cognition is strongly constrained by one's surrounding environment and physical state (Atkinson, 2010). Participants performed a Go-Nogo task in which they heard action verbs before manipulating virtual objects using real and varied actions, in an ecologically valid VR environment. We used time-frequency to measure motor activation, as evidenced by mu and beta ERD (Niccolai et al., 2014; Pfurtscheller & Lopes da Silva, 1999), during verb processing and prior to movement proper. We also examined ERP language-related components. We hypothesized that if motor activation proved to be contingent on language processing, this would go in the direction of embodied studies demonstrating that action language directly produces motor activation (Klepp, van Dijk, Niccolai, Schnitzler & Biermann-Ruben, 2019). Recording EEG in a VR environment is particularly challenging from a technical viewpoint, given the possibility of crosstalk between systems and physical movement (Török et al., 2014) as well as the need for precise synchronization between stimuli presentation and electrophysiological measures (Chapter 5, Tromp et al., 2018). Besides testing the hypothesis that motor activation is involved in linguistic retrieval, we aimed to provide proof of concept of the combination of EEG and VR.

1.6.4 The neural correlates of embodied L2 learning. Does embodied L2 verb learning affect representation and retention?

Chapter 6 presents a submitted registered report that describes the final work of this dissertation. In it, we once again combined EEG and VR, only this time to explore the neural correlates of embodied learning. This project aimed to extend the findings of our previous studies to a natural language learning environment. As in the experiment described in chapter 2, we will compare semantic processing and motor activation pre and post-training.

Participants' sensory-motor systems will be stimulated during an action verb learning task to test whether this reinforces sensory-motor representations of the new lexicon. Absolute beginners will learn an L2 lexicon of action verbs using a head-mounted VR display and perform physical actions such as picking up an object and throwing it right after hearing and repeating the verb "throw", for instance, or, in the control condition, simply pointing to the virtual object. Both pre and post-training, learners' knowledge of the semantic meaning of the new verbs will be measured behaviorally and using EEG. Mu and beta suppression will also be measured while participants listen to the new L2 verbs, both pre and post-training. We expect mu suppression during auditory verb processing to vary post-training as a function of learning condition. We hypothesize that representations of verbs learned in the embodied, specific action, condition will carry a stronger motor trace and hence produce greater mu suppression than verbs leaned in the control condition. We also predict that the specific action condition will lead to better learning outcomes post-training compared to the control condition, due to richer and more distributed representations.

Chapter 2

Cross-linguistic gender congruency effects during lexical access in novice L2 learners: evidence from EEG

Zappa, A., Mestre, D., Pergandi, J.-M.,& Frenck-Mestre, C., submitted to *Language, Cognition and Neuroscience*

Herein we present electrophysiological evidence of extremely rapid learning of new labels in an L2 (Brazilian Portuguese) for existing concepts, via computerized games. Interestingly, the effect was largely constrained by grammatical gender congruency. We recorded event-related potentials (ERPs) both prior to exposure to the second language (L2) and following a 4-day training session. Results showed rapid changes in cortical activity associated with learning. Prior to exposure, no modulation of the N400 component was found as a function of the correct match vs mismatch of audio presentation of words and their associated images. Post-training, a large N400 effect was seen for mismatch trials compared to correctly matched audio-visual trials. However, these results were only obtained for trials on which the L2 words shared grammatical gender in the learners' L1 (French). For trials on which the L2 words had the opposite gender in French, no N400 effect was found post-training. In contrast, behavioral results showed that all L2 words were learned equally as well, independent of gender con-

gruency across Portuguese and French. For control participants who were native speakers of Portuguese, a clear N400 effect was found for mismatch compared to match trials, which was independent of gender congruency. The results demonstrate that grammatical gender overlap in the L1 and L2 influences lexical activation during the initial stages of establishing a new L2 lexicon.

2.1 Introduction

In Brazilian Portuguese, a mouse, no matter which biological sex, is grammatically masculine (o *masc* camundongo *masc*) whereas a cockroach is feminine (a *fem* barata *fem*). The opposite is true in standard French, with grammatically feminine mice (la *fem* souris *fem*) and masculine cockroaches (le *masc* cafard *masc*). This arbitrary assignment of grammatical gender is even more apparent for inanimate objects, with opposite gender assignment for a screw, a broom and chalk across Portuguese and French, despite both languages being derived from Latin. The present study examined how gender congruency across languages might affect both the acquisition and online processing of a second language (L2) in novice adult learners. Although numerous online studies have provided evidence that speakers of gendered languages are sensitive to gender congruency across languages, during both L2 comprehension and production (cf. Sá-Leite, Fraga & Comesaña, 2019, for a meta analysis), these studies have almost exclusively examined learners who had extensive experience with the L2. We propose a novel approach to this question by starting from the initial stages of exposure to the L2. In addition, we provide both behavioral and electrophysiological measures of performance, which revealed different patterns of the effect of gender congruency.

Numerous psycholinguistic studies have examined the effect of gender congruency across languages (GCE) in bilinguals of varying proficiency levels and ages. However, the theoretical question addressed in these studies varies. Several bilingual studies have been conducted in the framework of speech production models, with the question being where gender is stored and when it is accessed (Caramazza, 1997; Schiller & Caramazza, 2006, vs Levelt, Roelofs & Meyer, 1999; Nickels, Biedermann, Fieder & Schiller, 2015). As concerns comprehension, bilingual studies examining the GCE have looked at both the interactive nature of bilingual lexical access (Morales et al., 2016) and at late bilinguals' ability to use grammatical gender to predict upcoming elements (Hopp & Lemmerth, 2018; Lemmerth & Hopp, 2019). The latter have addressed processing at the lexical level. However, only a handful of comprehension studies have also investigated the influence of gender congruency on syntactic processing (Foucart & Frenck-Mestre, 2011; Sabourin, Stowe & de Haan, 2006). We shall address these topics in turn.

Cross-linguistic GCEs have been examined at the lexical level during production, in naming (Bordag, 2004; Bordag & Pechmann, 2007; Costa, Kovacic, Franck & Caramazza, 2003; Lemhöfer, Spalek & Schriefers, 2008; Morales, Paolieri & Bajo, 2011) and translation (Bordag & Pechmann, 2008; Salamoura & Williams, 2007), as well as in comprehension (Lemhöfer et al., 2008). The pattern of results across studies is both complex and inconsistent. In two independent experiments with German (L1)-Dutch (L2) late bilinguals, Lemhöfer et al. (2008) examined the effects of cognate status and gender congruency on lexical decision times and naming latencies in the L2. They reported robust effects of both factors in both tasks, with no interaction effects. Participants showed faster lexical decision times and naming latencies for L2 Dutch words that shared gender in German and for cognates (cf. Costa, Caramazza & Sebastian-Galles, 2000, for a discussion of cognate effects). In addition, in the naming task, no effect of syntactic structure was found, such that GCEs were reported independently of whether participants produced determiner phrases or bare nouns. Similar results were reported by Bordag & Pechman (2007) for relatively inexperienced Czech (L1)-German (L2) late learners; they found a GCE in two experiments irrespective of whether participants named bare nouns, or nouns preceded by gender-marked adjectives in the L2. Both studies argue in favor of an interactive activation model of language processing (Dell, 1986) in which both the L1 gender and phonological form of lexical entries influence L2 processing, and according to which grammatical gender is not stored as an independent feature at the lemma level (cf. Levelt et al., 1999).

In contrast, Costa and colleagues failed to find a GCE across five independent experiments in which participants produced NPs in their L2, even when gender-marked determiners were used (Costa et al., 2003). This was true irrespective of whether the bilinguals' two languages had similar gender systems. The authors argued that while semantic representations are shared across languages and commonly activated by lexical entries of either language, the specific grammatical features of a lexical entry such as its gender are an inherent property of that entry. Hence, these features would not be shared across languages. Costa et al. (2003) noted nonetheless that they tested highly proficient bilinguals and suggested that less proficient L2 speakers might show greater interaction between their two gender systems, as was indeed reported by subsequent studies (Bordag & Pechman, 2007; Lemhöfer et al., 2008). In translation tasks, the results are also mixed. Salamoura & Williams (2007) reported a GCE when proficient Greek-German bilinguals translated gender-marked adjectives along with the nouns, but not for bare nouns. The authors argued that gender retrieval occurs only when gender concord must be computed, i.e. within the determiner phrase, in line with certain monolingual models of production (Caramazza, 1997). This contrasts, however, with the results found across three experiments for Czech-German late learners (Bordag & Pechmann, experiment 3, 2008), in which no GCE was found when participants translated either bare nouns or DPs into the L2. It is of interest to note that the absence of a GCE was reported in translation for the same materials and participant population that produced robust GCE in production (Bordag & Pechmann, 2007).

The effect of gender congruency across languages has also been examined during online comprehension. Several studies have approached this topic in the framework of whether bilinguals can use grammatical gender to predict upcoming elements in their L2 (Hopp & Lemmerth, 2018; Lemmerth & Hopp, 2019; Morales et al., 2016). The results from two visual world paradigm studies with Russian-German bilinguals, which used the same design and materials, failed to produce statistically conclusive evidence that gender congruency plays a significant role in the ability of either adults (Hopp & Lemmerth, 2018) or children (Lemmerth & Hopp, 2019) to process gender online in the L2. Another visual world paradigm study, conducted with proficient Italian-Spanish bilingual adults, showed interference from a distracter image when its gender was incongruent with the L1 equivalent, thereby suggesting gender-induced competition (Morales et al., 2016). However, the effects were not significant until after the onset of the target noun, indicating that co-activation of gender across languages during comprehension may not occur until a certain amount of information has been processed.

Rodríguez-Fornells & Münte (2016) recorded ERPs (event-related potentials) in a Go-Nogo paradigm to probe the effect of grammatical gender congruency (and language switching) across German and Spanish in fluent bilinguals. Compared to monolingual controls, bilinguals showed greater negativity for incongruent compared to congruent gender trials. While this result may indicate the automatic activation of gender across languages, there are several caveats. First, the task explicitly required participants to retrieve grammatical gender and both languages were actively recruited. Second, participants could potentially predict the incongruent gender trials based on the structure of the experiment. Last, gender incongruence elicited a late ERP component (P600/LPC). Hence, these results do not provide clear evidence for the automatic, early retrieval of gender. On the other hand, in a semantic categorization task conducted exclusively in English with Spanish-English late bilinguals and monolingual controls, Boutonnet, Athanasopoulos and Thierry (2012) reported a late negative ERP component (starting at roughly 400 ms) for trials that did not share gender in Spanish (the L1) with the two preceding items. This effect was specific to the bilingual group. No effects were found for the behavioral measures, although the absence of an effect for behavioral measures in the presence of electrophysiological evidence is rather common. The authors argued for the automatic activation of L1 gender in speakers of gendered languages, even when processing exclusively the L2 and in a non-gendered language such as English.

Gender congruency effects have also been examined at the syntactic level during sentence processing, using ERPs. Several monolingual studies have shown that gender agreement violations in sentential context systematically elicit the P600 component (Barber & Carreiras, 2005; Frenck-Mestre, 2005; Gunter, Friederici & Schriefers, 2000; Hagoort, 2003; Popov & Bastiaanse, 2018; Popov, Miceli, Ćurčić-Blake & Bastiaanse, 2020; Wicha, Moreno & Kutas, 2004) and can also produce a LAN (Barber & Carreiras, 2005; Gunter et al., 2000; Popov et al., 2020; for a discussion of the LAN cf. Tanner, 2015 vs Molinaro, Barber, Caffarra & Carreiras, 2015). The same is true during L2 sentential processing for L2 learners whose L1 has grammatical gender (Foucart & Frenck-Mestre, 2011; Sabourin et al., 2006) but also for those whose L1 does not (Foucart & Frenck-Mestre, 2012; Dowens, Guo, Guo, Barber & Carreiras, 2011; Morgan-Short, Sanz, Steinhauer & Ullman, 2010; Tokowicz & MacWhinney, 2005). Moreover, P600 amplitude for gender concord violations in an L2 is contingent on proficiency and age of acquisition (Nichols & Joanisse, 2016). ERP studies that focused on cross-language gender congruency have provided evidence that the overlap of both syntactic rules, as concerns gender concord (Foucart & Frenck-Mestre, 2011; Sabourin & Stowe, 2008), and lexical gender across languages (Foucart & Frenck-Mestre, 2011) affect whether gender concord violations in the L2 elicit an electrophysiological response, the type of response (Foucart & Frenck-Mestre, 2012) and its magnitude.

The body of studies cited above has examined L2 gender processing and gender congruency in participants who had several years of experience with and exposure to the L2. Various authors have used learning paradigms with either an artificial language or miniature versions of natural languages to explore how different factors affect gender acquisition after short training periods. Arnon and Ramscar (2012) used an artificial language to test whether acquiring the gender and new lexical labels of known concepts was affected by the sequence of explicit training. Participants who first learned novel words preceded by their gender-marked article within an auditory sentential context, followed by paired associate learning with bare nouns, had better learning outcomes, both for gender assignment and noun labels, than those who learned in the opposite order. The authors argued that learning new lexical labels for concepts first via bare nouns blocked the later acquisition of gender assignment in sentential context due to the redundancy of gender in relation to meaning. Brooks and Kempe (2013) examined the implicit learning of nominal gender agreement and case marking in a subset of Russian following a 6-day training session. Results showed that while learners relied on metalinguistic knowledge to acquire case marking, for gender concord they relied on consistent and transparent morphological cues (feminine being systematically indicated by a final vowel on the noun and agreeing adjective, and masculine by a final consonant) and knowledge of nominal morphology in another known L2. Indeed, the best predictor of acquiring Russian gender agreement was whether the learners had already acquired a Latin language with the same rule for feminine gender. Morgan-Short and colleagues (2010) compared the processing of gender concord in early learning stages, after implicit versus explicit training in an artificial language. During early stages of acquisition, the ERP signature to gender concord violations differed according to the type of training. However, both implicit and explicit learning groups ultimately attained similar levels of proficiency and exhibited similar patterns of cortical response to gender concord violations at the final stage of acquisition.

To our knowledge, no studies have measured the effect of cross-language gender congruency during the early stages of L2 lexical acquisition in a natural language. The current study aimed to fill this void by examining how gender congruency might influence processing in an L2 from the very initial stages of learning. We examined French L1 speakers' acquisition of Brazilian Portuguese via interactive computer-games, in which L2 Portuguese was presented auditorily in both full sentences and segmented format and in which grammatical gender within determiner phrases was taught implicitly. Both French and Portuguese have two classes of grammatical gender (masculine and feminine) and require gender concord within the determiner phrase. Whereas French uses the singular definite article "le" *masc* [lə] to mark the masculine gender and Portuguese uses "o" *masc* (realized as /o/ or /u/), in both languages the singular feminine definite article carries the final phoneme [a] (French: la *fem* [la], Portuguese: a *fem* [a]). In addition, in Portuguese, the vowel of the definite determiner is generally consistent with the final vowel of the noun (e.g."a faca" [the knife] and "o garfo" [the fork]). It is therefore possible that, even without formal instruction concerning the gender of the Portuguese nouns or the determiner system, French native speakers are able to extract this information from the phonological word forms (Brooks & Kempe, 2013; Denhovska & Serratrice, 2017).

We created four computer-games to teach French native speakers a small lexicon in Brazilian Portuguese. All auditory materials were presented exclusively in Brazilian Portuguese. The games involved both full sentences and individual lexical items. We manipulated gender congruency such that the nouns were either gender congruent or incongruent across the learners' L1 and L2 (cf. Table 1). No instruction was provided concerning grammatical gender; however, nouns were always preceded by the singular definite determiner, i.e. the overtly marked vowel (e.g. a *fem* faca *fem* [the knife], o *masc* garfo *masc* [the fork]). Participants with no prior knowledge of (Brazilian) Portuguese took part in a 4-day training program including 3 training days during which they learned a vocabulary of 12 lexical items comprised of 3 verbs and 4 nouns each day via the games and 1 review day. ERPs were recorded both pre and post-training in a match-mismatch paradigm in which auditory nouns were paired with visual images that either depicted the noun (match) or another learned noun (mismatch). Thus, we were able to follow progression from zero knowledge to the recognition of newly learned L2 phonological word forms.

We hypothesized that learners should be able to fully acquire the L2 vocabulary. Performance was measured both by their accuracy scores pre and post-training and, crucially, by the change in electrophysiological response pre to post-training. Concerning the latter, we expected variation in the N400 effect, whereby prior to training match and mismatch trials should not differ in the N400 response, but post-training mismatch trials should evoke an increased N400 compared to match trials due to difficulties in lexical processing. In relation to the congruency of grammatical gender across the L1 and L2, we hypothesized that it should affect learning as indexed by a post-training N400 effect if indeed grammatical gender is automatically activated for speakers of gendered languages (Boutonnet et al., 2012; Dahan, Swingley, Tanenhaus & Magnuson, 2000; Lew-Williams & Fernald, 2007) and if inhibition results from gender inconsistency across languages (Morales et al., 2016; Rodríguez-Fornells & Münte, 2016).

2.2 Methods

2.2.1 Participants

Eighteen right-handed French native speakers, enrolled as undergraduate students at Aix-Marseille Université (10 women), aged 20 to 26 years old (M = 22.8 years, SD = 2.4), who had no knowledge of Brazilian Portuguese, and 18 right-handed native speakers of Brazilian Portuguese aged 22 to 28 (M = 25.3, SD = 3.4) were recruited for the study. One French participant's data was excluded due to displaying knowledge of the L2 vocabulary (an N400 effect for mismatched pairs) prior to training. Participants had no history of neurological insult and received monetary compensation for their participation. All participants gave their written informed consent prior to the experiment and were debriefed about its purpose at its end. The study was approved by the local university ethics committee.

2.2.2 Materials

Thirty-six concrete nouns and 9 transitive verbs were presented orally in Brazilian Portuguese, in sentences and in isolation, across 3 training sessions. The items were selected based on their ease of imageability, non-transparency with the learners' L1 (French) and other Latin languages (Italian, Spanish) and English, and congruency of grammatical gender across Portuguese and French. Half of the nouns had the same gender across languages ([a *fem* panela *fem*] / la *fem* casserole *fem* "the pot") and the other half had the opposite gender ([a *fem* faca *fem*] / le *masc* couteau *masc* "the knife"). Nouns were systematically preceded by the definite determiner (/a/*fem* or /o/*masc*). Nouns that shared gender (SG) and those that had opposite gender (OG) across French and Portuguese were equated across numerous lexical variables: printed mean frequency per million in French (19.71(21.19)) vs (18.04(16.05)) for SG and OG, respectively (New, Pallier, Brysbaert & Ferrand, 2004), mean number of letters SG(6.17(1.70)) vs OG(5.53(1.62)), mean number of phonemes SG(5.50(1.62) vs OG(5.47(1.51)) and mean number of syllables SG(2.58 (0.79)) vs OG(2.47(0.80)) and grammatical gender SG(8*fem*/4*masc*) vs OG(8*fem*/7*masc*).

Each of the 9 verbs was paired with 4 different nouns to create 36 declarative sentences in canonical SVO order preceded by a lead-in phrase (e.g. esfregar/scrub [Ele está esfregando a janela/a lareira com a escova/o trapo] "He is scrubbing the window/fireplace

Table 2.1: Pre-post for L2

	L2fem	L2masc
L1fem	a saia(la jupe)	o cachimbo(la pipe)
L1masc	a vassoura(le balai)	o casaco(le manteau)

with the brush/rag"). Three additional partially transparent verbs (e.g. [pintar], "peindre", "to paint") and 12 additional partially transparent nouns (e.g. [esponja], "éponge", "sponge") were selected to familiarize participants with the games and for electroencephalography (EEG) testing. All auditory materials were recorded by a native Brazilian female speaker at 48 kHz (32-bit float) in a professional sound booth, in a single session. The onsets of each word within auditory sentences were detected automatically using SPPAS (Bigi, 2015) and manually verified using PRAAT (Boersma & Weenink, 2018). The materials were spliced into individual syntactic units (pronoun + copula, determiner + noun, lexical verb) and individual sentences using Audacity 2.2.1 software.

A subset of the materials was selected for pre and post-training tests in a match-mismatch paradigm using EEG. Thirty of the auditory nouns, half with the same gender and half with the opposite gender across French and Portuguese (Table 2.1), were paired with 30 line drawings selected from the Snodgrass & Vanderwart (1980) and Alario & Ferrand (1999) standardized image databases (with the exception of two images which were taken from line-drawing internet databases). These images were different from those used for the computerized games (described below), which were selected from internet databases. Each auditory noun was presented twice, once paired with the correct line drawing (match) and once with a line drawing that corresponded to another (to be) learned noun (mismatch). Grammatical gender was neutralized in mismatch trials such that the noun depicted in the image had the same gender in Brazilian Portuguese as the word presented orally (e.g. the image of a knife [a faca] was presented with the auditory noun [a lareira] "the fireplace"). Three pseudo-randomized lists were created for pre-training EEG testing including 30 Match pairs and 30 Mismatch pairs. Three other pseudo-randomized lists were created for post-training EEG testing, which included the same 30 Match and 30 Mismatch pairs and an additional 30 semantically related and 30 semantically unrelated pairs (data reported elsewhere). Participants saw different lists at pre and post-training testing.

2.2.3 Procedure

Games

Four computerized games were created in collaboration with the Mediterranean Virtual Reality Center (CRVM). All 4 games involved the auditory presentation of materials in Brazilian Portuguese via headphones, accompanied by either static line drawings or animated GIFs on a flat screen. Participants' responses and playing behavior (mouse clicks, timing) were recorded throughout each game and feedback was provided. Trials that were not completed successfully were repeated at the end of each game. All participants played the 4 games in the same order and were required to successfully complete a given game prior to engaging in the next.

In all 4 games, participants initiated a trial by clicking on an audio button image to hear an auditory stimulus. In the first, "exposure", game, participants clicked to hear the 12 sentences, one at a time, which were presented simultaneously with an animated GIF of the action and accompanying objects. In the second, "segmentation", game, participants clicked to hear a sentence, which was accompanied by the visual presentation of 5 blank squares at the bottom and 3 in the centre of the computer screen. Participants clicked on any of the 5 bottom squares to display a static image and hear the audio file corresponding to it (4 nouns and 1 verb were depicted). Participants had to recreate the auditory sentence by clicking and dragging the 3 correct syntactic elements in order (verb, NP1, NP2) to the centre of the screen. Upon correct completion, an animated GIF played along with the auditory sentence. In the third, "verb identity", game, participants clicked to hear a verb and saw three different animated GIFs in the center of the screen, the task being to select the animation that corresponded to the audio file. In the final, "memory", game, participants clicked to hear a sentence, then had to find the pair of cards, among 8 presented face down on the screen, that matched the auditory sentence by clicking on the cards individually (a card reverted to blank when another card was selected). The vocabulary depicted across the 8 cards involved a single verb and 4 nouns. When the correct pair was selected, the auditory sentence was replayed along with an animated GIF. Across games 2, 3 and 4, participants were allowed to click on a given item (audio button or card) a maximum of 3 times and time-out was 30 seconds on any given trial. They were encouraged to repeat the materials out loud while playing. Participants were not permitted to take notes during training sessions and were asked not to review what they had learned between sessions.

Training

Participants learned how to play the 4 computer-games, prior to actual training, using a miniature auditory vocabulary. This initiation to the games took place directly after the first EEG session. During the initiation and subsequent training, participants were comfortably seated in a sound-attenuated room where they played the games on a 15-inch-screen laptop computer while wearing headphones. Training occurred over 4 consecutive days. On each of the first three days, participants were exposed to 3 verbs and 12 nouns, comprised in 12 auditory sentences, via the 4 games, with the vocabulary repeated across the games. Each session lasted roughly 30 minutes, with each game taking 5-10 minutes. The fourth day consisted of a 40-minute review, where participants played all the games with the entire new lexicon (9 verbs and 36 nouns).

2.2.4 EEG Match-Mismatch task

A trial began with the presentation of a centered fixation cross for 250 ms that was replaced by a centered black and white line drawing for 1 second followed by an auditory word presented over speakers. At the offset of the auditory word, a visual “yes/no” prompt was presented and participants were requested to judge whether the auditory word matched the visual image or not on a button box. A visual blink prompt was presented for 2s following the response. During the pre-training session, 2 blocks of 30 trials, with 15 “match” and 15 “mismatch” pairs in each, were presented, preceded by 3 warm-up pairs. During the post-training session, an additional 30 trials were presented per block, comprising 15 semantically related and 15 unrelated pairs (data reported elsewhere), for a total of 60 trials per block. Short rest periods were allowed between blocks. Participants were asked to remain still and to blink at the prompt. Behavioral responses to the questions were recorded. The session lasted roughly 60 minutes.

2.2.5 ERP Data acquisition and analysis

EEG activity was recorded continuously from 64 scalp locations over frontal, temporal, central, posterior temporal, parietal and occipital areas of the left and right hemispheres and midline. Individual electrodes were adjusted to a stable offset lower than 20µV. EEG data were sampled online at 512 Hz. Blinks and horizontal eye-movements were monitored by means of electrodes placed beneath the left eye and at the outer canthus of the right eye. Electrodes were placed on the left and right mastoids for referencing offline. Periods spanning from -100 pre-stimulus onset to 1100 msec post stimulus onset were used post-recording for analyses. A low pass digital filter of 30 Hz was applied post-recording. Trials contaminated by ocular-motor or muscular artifacts were excluded using automated routines that were manually checked. The percentage of trails retained for analyses was 88% for the same gender condition and 87% for the opposite gender condition.

2.3 Results

2.3.1 Behavioral results

The d-prime scores are presented in Figure 2.1 for the L2 learners at the pre-training session and the post-training session, and for the Brazilian control group. They modeled using linear mixed effect regressions, with the LmerTest package (Kuznetsova, Brockhoff & Christensen, 2017) implemented in R (R Core Team, 2017). We first modeled the data for L2 learners, pre and post-training including the sum coded fixed factors Training session (Pre vs Post), Gender (Same vs Opposite) and their interaction, with random intercept for Participant. The model revealed an effect of Training session (β= 2.44, se = 0.205, t = 11.92, $p<.001$) and no other effects (cf. Table 2.2). Participants' d-prime scores increased significantly from pre to post-test and this did not depend on the gender congruency of trials.

We subsequently modeled the dprime scores for L2 learners post-training and the native Brazilian control group. The effect of Group was significant (β= 1.55, se = 0.239, t = 6.50, $p<.001$), due to Brazilian controls showing higher d prime scores. No other effects were significant (cf. Table 2.3).

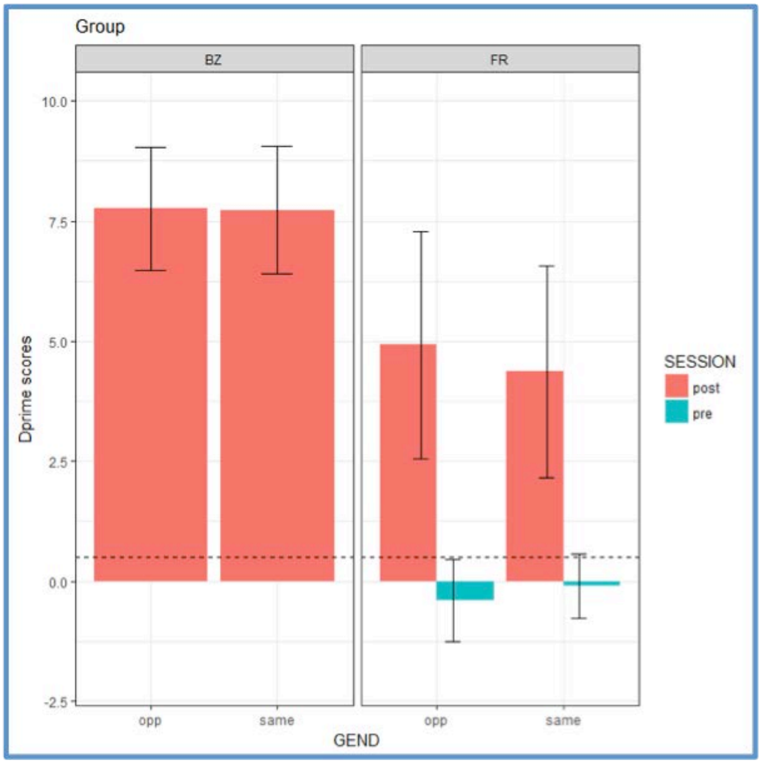


Figure 2.1: *Dprime scores for Brazilian controls and for L2 learners at pre and post-test sessions*

2.3.2 ERP analysis

The ERP data were modeled using linear mixed effect regressions, with the *LmerTest* package (Kuznetsova et al., 2017) implemented in R (R Core Team, 2017) for the mean voltage amplitudes in the N400 time window, calculated 300-600 msec post onset of the auditory noun for correct trials. Data were trimmed in R to remove outliers (1% of the data were excluded). The data for L2 learners were first modeled for the pre-training and post-training session. To compare learners to native speakers, data from the post-training session were then modeled for both groups. Models were performed independently over midline sites (Fz, FCz Cz, CPz, Pz), frontal-central sites (FC1, FC3, FC5, FC2, FC4, FC6, C1, C3, C5, C2, C4, C6) and centro-parietal sites (CP1, CP3, CP5, CP2, CP4, CP6, P1, P3, P5, P2, P4, P6). Below we report the results from the maximal random-effects structure (Barr, Levy, Scheepers & Tily, 2013).

Pre vs Post-training: L2 learners

To examine the effect of learning, we ran a first model, summarized in Table 2.4, which included the sum-coded fixed factors Training session (Pre vs Post), Gender (Same vs Opposite), Condition (Match vs Mismatch) and their interactions, with random intercepts for Participant and Item. Condition included a random slope for Participant and for Item. The model revealed a three-way interaction of Condition:Gender:Session at all electrode sites. The data were modeled independently thereafter for each training session. Pre-training, no effects were found for any factor at any ROI (cf. Table 2.5). Figures 2.2 and 2.3 show the mean Match-Mismatch ERP response for Same and Opposite gender conditions, respectively, for illustrative purposes.

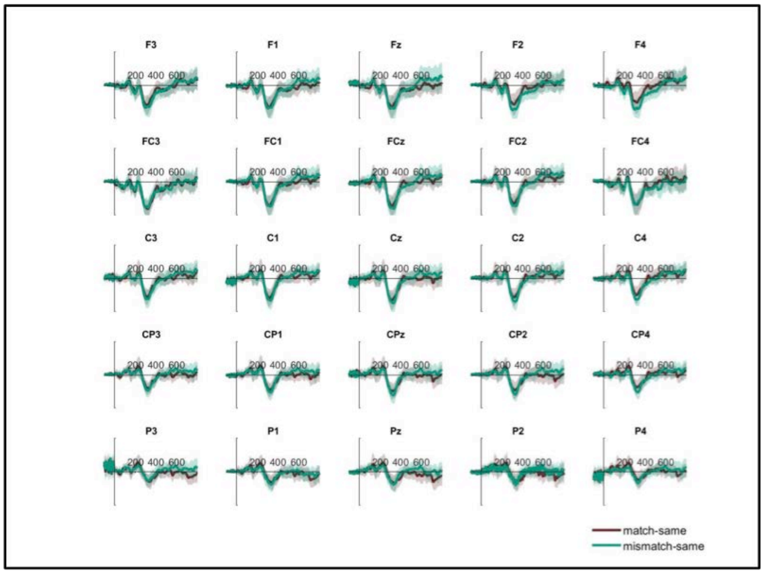


Figure 2.2: *Mean voltage ERPs (and SD) in the pre-training session for L2 learners as a function of Condition (Match vs Mismatch pairs) for nouns with same gender across languages*

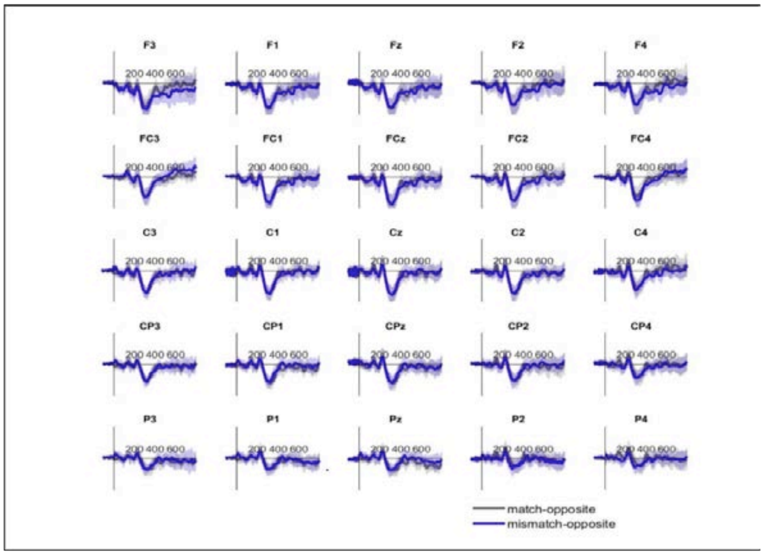


Figure 2.3: *Mean voltage ERPs (and SD) in the pre-training session for L2 learners as a function of Condition (Match vs Mismatch pairs) for nouns with opposite gender across languages*

Post-training, sum coded models performed independently at all 3 ROI revealed a significant interaction effect of Condition:Gender at all sites (cf. Table 2.6). Models of simple effects (cf. Tables 2.7 and 2.8) revealed a significant effect of Condition for same gender trials at all ROI (midline: β= -3.40, se = 0.770, t = -4.42, $p<.001$; frontal-central: β= -2.60, se = 0.682, t = -3.82, $p<.002$; central-parietal: β= -2.91, se = 0.577, t = -5.05, $p<.001$) but no effect of Condition for opposite gender trials (midline: β= -0.27, se = 0.757, t = -0.36 $p<0.73$; frontal-central: β= -0.52, se = 0.644, t = -0.81, $p<.43$; central-parietal: β= -0.59, se = 0.613, t = -0.97, $p<.35$). These effects are shown in Figures 2.4 and 2.5.

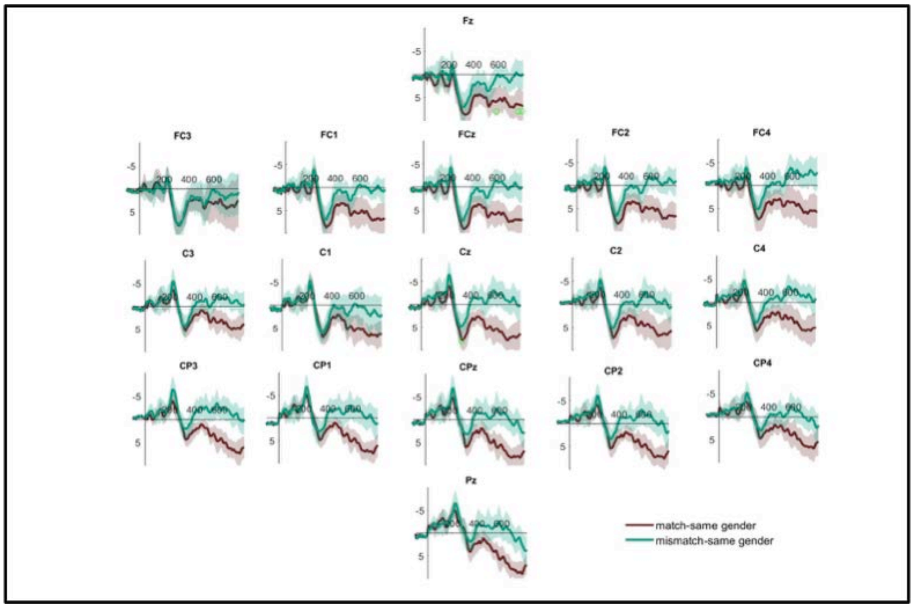


Figure 2.4: *Mean voltage ERPs (and SD) in the post-training session for L2 learners as a function of Condition (Match vs Mismatch) for nouns with same gender across languages*

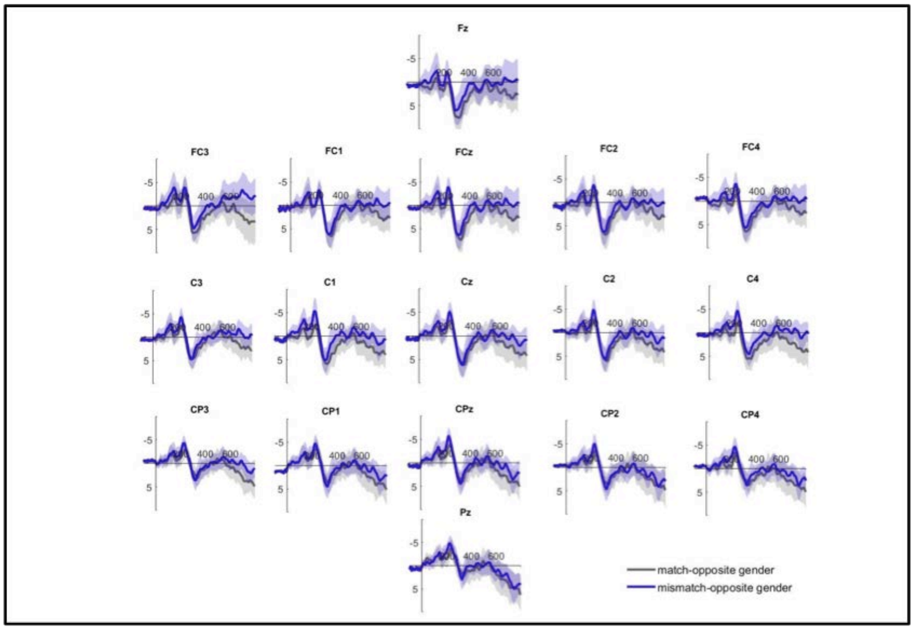


Figure 2.5: *Mean voltage ERPs (and SD) in the post-training session for L2 learners as a function of Condition (Match vs Mismatch) for nouns with opposite gender across languages*

In addition, for Match trials there was an effect of Gender at all three ROI (midline: β= 2.08, se = 0.422, t = 4.92, p<.001; frontal-central: β= 1.69, se = 0.584, t = 2.89, p<.01; central-parietal: β= 1.27, se = 0.257, t = 4.96, p<.001), while for Mismatch trials the effect of Gender was only present at midline (midline: β= -0.91, se = 0.434, t = -2.09, p<.04; frontal-central: β= -0.38, se = 0.805, t = -0.48, p<0.64; central-parietal: β= 1.06, se = 0.89, t = -1.19, p<.0.25). These effects are depicted in Figures 2.6 and 2.7.

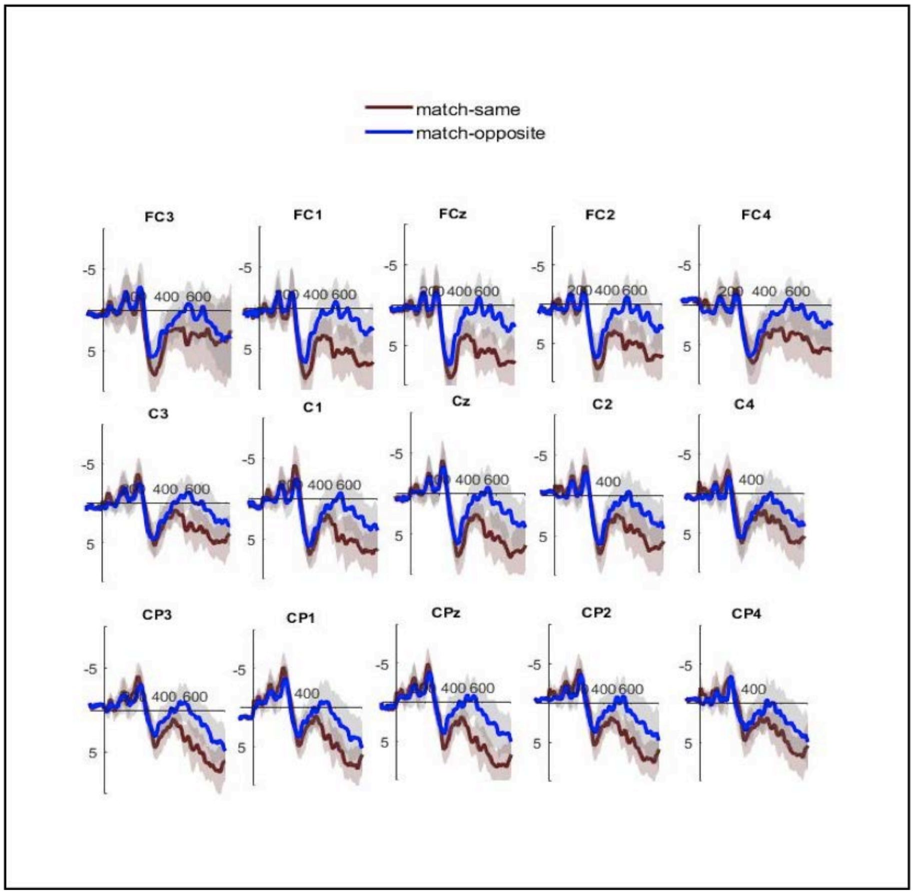


Figure 2.6: *Mean voltage ERPs (and SD) in the post-training session for L2 learners as a function of Gender across languages (same vs opposite) for Match trials*

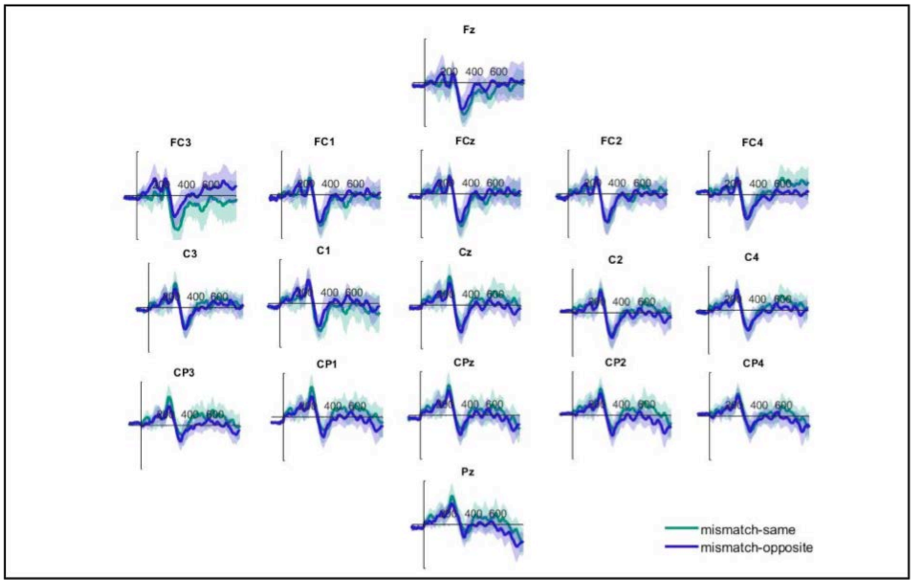


Figure 2.7: *Mean voltage ERPs (and SD) in the post-training session for L2 learners as a function of Gender across languages (same vs opposite) for Mismatch trials*

Post-training: Native vs L2 learners

The first model, summarized in Table 2.9, included the sum-coded factors Group (Native vs L2), Gender (Same vs Opposite), Condition (Match vs Mismatch) and their interactions. Condition included a random slope for Participant and for Item. The model revealed an interaction of Condition:Gender:Group at all sites. The data were modeled independently thereafter for the Brazilian control group, using the same model structure as above without the fixed factor Group. For native speakers, the only significant effect was of Condition, with greater mean N400 amplitude for mismatch than match trials at all sites (see Table 2.10).

2.4 Tables

Table 2.2: **Model output for d-prime scores for L2 learners at pre and post-test**
lmer(Dprime~(1|SUBJECT) + SESSION.sum + GEND.sum + SESSION.sum:GENDER.sum

Fixed effects: Midline Sites	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	2.19790	0.21219	10.358	1.68e-08 ***
SESSION.sum1	2.44481	0.20506	11.922	5.91e-16 ***
GEND.sum1	0.06119	0.20506	0.298	0.767
SESSION.sum1:GEND.sum1	0.21407	0.20506	1.044	0.302

Table 2.3: **Model output for d-prime scores for L1 Brazilians and and L2 learners at post-test**
lmer(Dprime~(1|SUBJECT) + GROUP.sum + GEND.sum + GROUP.sum:GENDER.sum)[.5cm]

Fixed effects: Midline Sites	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	6.1925	0.2386	25.956	2e-16 ***
GROUP.sum1	1.5498	0.2386	6.496	2.26e-07 ***
GEND.sum1	0.1455	0.2020	0.720	0.476
GROUP.sum1:GEND.sum1	-0.1298	0.2020	-0.642	0.525

Table 2.4: **Summary of effects for L2 learners for the 3 ROI pre and post-training**
lmer(MVC ~(1+COND.sum |SUBJECTS) + (1+COND.sum |ITEM) + COND.sum + GEND.sum + EXPE.sum + COND.sum:GEND.sum + GEND.sum:EXPE.sum + COND.sum:EXPE.sum + GEND.sum:EXPE.sum:COND.sum)

Fixed effects: Midline Sites	Estimate	Std.Error	t value	Pr(>|t|)
Intercept)	1.24420	0.53291	2.335	0.0328*
COND.sum1	00.43901	0.20430	2.149	0.0442*
GEND.sum1	-0.19069	0.10448	-1.825	0.0680
EXPE.sum1	0.04772	0.10497	0.455	0.6494
COND.sum1:GEND.sum1	-0.30688	0.13645	-2.249	0.0338*
GEND.sum1:EXPE.sum1	-0.10439	0.10450	-0.999	0.3179
COND.sum1:EXPE.sum1	0.47717	0.10474	4.556	5.31e-06 ***
COND.sum1:GEND.sum1:EXPE.sum1	-0.42132	0.10454	-4.030	5.63e-05 ***
Fixed effects : Frontal Central Sites	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	9.191e-01	4.801e-01	1.915	0.07348
COND.sum1	3.745e-01	1.859e-01	2.015	0.06091
GEND.sum1	-2.213e-01	6.605e-02	-3.350	0.00255 **
EXPE.sum1	5.209e-02	6.349e-02	0.820	0.41200
COND.sum1:GEND.sum1	-2.581e-01	6.322e-02	-4.082	4.48e-05 ***
GEND.sum1:EXPE.sum1	-1.010e-01	6.325e-02	-1.597	0.11020
COND.sum1:EXPE.sum1	4.016e-01	6.344e-02	6.331	2.50e-10 ***
COND.sum1:GEND.sum1:EXPE.sum1	-2.057e-01	6.326e-02	-3.251	0.00115 **
Fixed effects : Central Parietal Sites	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	5.525e-01	4.365e-01	1.266	0.224
COND.sum1	4.020e-01	1.799e-01	2.234	0.040 *
GEND.sum1	-7.605e-02	6.285e-02	-1.210	0.226
EXPE.sum1	-8.894e-03	6.317e-02	-0.141	0.888
COND.sum1:GEND.sum1	-2.463e-01	6.285e-02	-3.919	8.92e-05***
GEND.sum1:EXPE.sum1	-3.943e-02	6.286e-02	-0.627	0.530
COND.sum1:EXPE.sum1	4.686e-01	6.312e-02	7.424	1.19e-13***
COND.sum1:GEND.sum1:EXPE.sum1	-2.634e-01	6.287e-02	-4.189	2.81e-05***

Table 2.5: **Summary of effects at the pre-training session for L2 learners for the 3 ROI**

lmer(MVC ~(1+COND.sum|SUBJECTS) + (1+CONDITEM) + COND.sum + GEND.sum + COND.sum:GEND.sum

Fixed effects: Midline Sites	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	1.20538	0.58093	2.075	0.0533
COND.sum1	-0.02400	0.22892	0.105	0.9177
GEND.sum1	-0.09297	0.17807	0.522	0.6061
COND.sum1:GEND.sum1	0.06651	0.14335	0.464	0.6427

Fixed effects: Frontal-central Sites	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	8.775e-01	5.270e-01	1.665	0.115
COND.sum1	7.535e-03	1.714e-01	0.044	0.965
GEND.sum1	-1.183e-01	8.777e-02	-1.347	0.178
COND.sum1:GEND.sum1	-8.305e-02	8.783e-02	-0.946	0.344

Fixed effects: Central Parietal Sites	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	0.58026	0.44970	1.290	0.215
COND.sum1	-0.05230	0.20715	-0.252	0.804
GEND.sum1	-0.03018	0.08599	-0.351	0.726
COND.sum1:GEND.sum1	-0.01232	0.08601	-0.143	0.886

Table 2.6: **Summary of effects at the post-training session for L2 learners for the 3 ROI**
lmerMVC ~(1.sum|SUBJECTS) + (1COND.sum |ITEM) COND.sum GEND.sum COND.sum:GEND.sum

Fixed effects: Midline Sites	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	0.4832	0.5397	1.117	0.28032
COND.sum1	0.8596	0.2227	3.859	0.00133**
GEND.sum1	-0.1175	0.0911	1.289	0.19730
COND.sum1:GEND.sum1	-0.5198	0.0911	-5.705	0.000000012 ***
Fixed effects : Frontal Central Sites	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	0.95444	0.53177	1.795	0.9117
COND.sum1	0.74904	0.24744	3.027	0.00779 **
GEND.sum1	-0.32470	0.10393	-3.125	0.00433 **
COND.sum1:GEND.sum1	-0.48185	0.09068	-5.314	0.000000133 ***
Fixed effects: Central Parietal Sites	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	0.5397	0.4832	1.117	0.28032
COND.sum1	0.8596	0.2227	3.859	0.00133 **
GEND.sum1	-0.1175	0.0911	-1.289	0.19730
COND.sum1:GEND.sum1	-0.5198	0.0911	-5.705	0.000000012 ***

Table 2.7: **Simple effects of Condition, at the post-training session for L2 learners, for same and opposite gender trials**
lmerMVC ~(1+COND|SUBJECTS) + (1|ITEM)

Fixed effects: Midline Sites	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	3.2598	0.6633	4.915	0.000146***
CONDmis	-3.4028	0.7703	-4.417	0.000525***

Fixed effects : Frontal Central Sites	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	2.559	0.7118	3.596	0.00225
CONDmis	-2.6034	0.6816	-3.820	0.00151**

Fixed effects: Central Parietal Sites	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	2.0111	0.5935	3.388	0.003669 **
CONDmis	-2.9107	0.5770	-5.045	0.000127***

Fixed effects: Midline Sites	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	1.0730	0.8901	1.205	0.244
CONDmis	-0.2687	0.7569	-0.355	0.727

Fixed effects : Frontal Central Sites	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	0.8607	0.8055	1.069	0.301
CONDmis	-0.5241	0.6441	-0.814	0.427

Fixed effects: Central Parietal Sites	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	0.7543	0.7770	0.971	0.346
CONDmis	-0.5932	0.6129	-0.968	0.347

Table 2.8: **Simple effects of Gender, at the post-training session for L2 learnes, for match and mismatch trials**
lmerMVC ~(1+GEND|SUBJECTS) + (1|ITEM)

Fixed effects: Midline Sites	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	1.1568	0.7233	1.599	0.126
GENDsame	2.0753	0.4222	4.916	9.62e-07

Fixed effects : Frontal Central Sites	Estimate	Std.Error	t value	Pr(>|t |)
(Intercept)	0.8604	0.8033	1.071	0.3000
GENDsame	1.6908	0.5854	2.888	0.0107

Fixed effects: Central Parietal Sites	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	0.7621	0.6029	1.264	0.223
GENDsame	1.2745	0.2570	4.960	7.19e-07***

Fixed effects: Midline Sites	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	0.8051	0.5736	1.404	0.1754
GENDsame	-0.9082	0.4340	-2.093	0.0365*

Fixed effects : Frontal Central Sites	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	0.3356	0.5399	0.621	0.543
GENDsame	-0.3838	0.8047	-0.477	0.640

Fixed effects: Central Parietal Sites	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	0.1590	0.7017	0.227	0.824
GENDsame	1.0637	0.8916	-1.193	0.250

Table 2.9: **Summary of effects at the post-training session for both groups and for the three electrode regions**

lmer(MVC ~(1+COND.sum|PARTICIPANTS + (1+COND.sum |ITEM) + COND.sum:GEND.sum + GEND.sum:GROUP.sum + COND.sum:GROUP.sum+ GEND.sum:GROUP.sum:COND.sum)

Fixed effects: L2 and Control at Midline Sites	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	1.0468	0.4197	2.494	0.01758
COND.sum1	1.1548	0.2038	5.665	2.59e-06
GEND.sum1	0.1021	0.1255	0.814	0.42275
GROUP.sum1	0.2257	0.4148	0.544	0.59009
COND.sum1:GEND.sum1	0.3167	0.1127	-2.811	0.00626
GEND.sum1:GROUP.sum1	0.1904	0.1081	1.760	0.07838
COND.sum1:GROUP.sum1	0.2434	0.2014	1.209	0.23548
COND.sum1:GEND.sum1:GROUP.sum1	0.4298	0.1081	3.975	7.11e-05
Fixed effects: L2 and Control at frontal-central Sites	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	8.075e-01	4.017e-01	2.010	0.052510
COND.sum1	9.765e-01	1.872e-01	5.218	9.42e-06***
GEND.sum1	3.382e-01	7.398e-02	4.572	0.000102***
GROUP.sum1	1.450e-01	4.003e-01	0.362	0.719407
COND.sum1:GEND.sum1	2.782e-01	6.611e-02	-4.208	2.80e-05***
GEND.sum1:GROUP.sum1	-1.400e-02	6.585e-02	-0.213	0.831647
COND.sum1:GROUP.sum1	2.248e-01	1.871e-01	1.202	0.237987
COND.sum1:GEND.sumsame:GROUP.sum1	2.039e-01	6.585e-02	3.096	0.001966**
Fixed effects: L2 and Control at central-parietal Sites	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	2.025e-01	3.656e-01	0.554	0.58341
COND.sum1	1.237e+00	2.034e-01	6.083	7.26e-07***
GEND.sum1	-1.375e-01	6.475e-02	-2.124	0.03368 *
GROUP.sum1	-3.365e-01	3.656e-01	-0.920	0.36398
COND.sum1:GEND.sum1	-2.873e-01	6.476e-02	-4.436	9.21e-06***
GEND.sum1:GROUP.sum1	-1.978e-02	6.475e-02	0.305	0.76001
COND.sum1:GROUP.sum1	3.761e-01	2.034e-01	1.849	0.07329
COND.sum1:GEND.sum1:GROUP.sum1	2.331e-01	6.476e-02	3.599	0.00032***

Table 2.10: **Summary of effects at the post-training session for Brazilian controls for the 3 ROI**
lmerMVC ~(1.sum|SUBJECTS) + (1.sum |ITEM) .sum .sum COND.sum:GEND.sum

Fixed effects: Midline Sites	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	0.82079	0.60273	1.362	0.190656
COND.sum1	1.39788	0.31593	4.425	0.000387 ***
GEND.sum1	0.08912	0.16977	0.525	0.603694
COND.sum1:GEND.sum1	0.11428	0.15576	0.734	0.463721

Fixed effects : Frontal Central Sites	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	0.66200	0.59535	1.112	0.281627
COND.sum1	1.20152	0.27855	4.313	0.000475 ***
GEND.sum1	-0.35212	0.09536	3.693	0.082317 COND.sum1:GEND.sum1
-0.07351	0.09536	0.771	0.440817	

Fixed effects: Central Parietal Sites	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	-0.13409	0.54457	-0.246	0.808454
COND.sum1	1.61456	0.33473	4.823	0.000157 ***
GEND.sum1	-0.15744	0.09181	-1.715	0.086405
COND.sum1:GEND.sum1	0.05374	0.09181	0.585	0.558323

2.5 Discussion

Our study revealed a clear cross-linguistic gender congruency effect (GCE), from the earliest stages of acquiring a second language. This was apparent in the electrophysiological trace of lexical access, as measured by the N400. L2 learners demonstrated an N400 effect for mismatched visual-auditory pairs post-training, but only for learned nouns that shared grammatical gender across their native (French) and newly acquired language (Portuguese). No modulation of the ERP response was found as a function of the match between auditory words and visual stimuli for nouns that had opposite gender across the two languages. In addition, the effect of gender congruency was visible in the N400 modulation for match trials, for which the N400 response was increased for nouns that had opposite gender across the L2 and the L1, compared to nouns that shared grammatical gender across languages. Hence, a GCE was clearly reflected in the automatic cortical response associated with lexical processing (Kutas & Federmeier, 2011). It is important to note that in the control group of native Brazilian Portuguese speakers, who on average did not speak French, only a robust N400 mismatch effect was found, which was independent of gender congruency. Otherwise stated, the GCE revealed by the N400 was specific to the L2 learners processing nouns in the newly learned language. Importantly, this effect was found following only 4 days of training using interactive computerized games.

Previous studies on the effect of cross-linguistic gender congruency have examined this question in populations that had several years of formal learning of the second language (Bordag, 2004; Bordag & Pechmann, 2007, 2008; Costa et al., 2003; Lemhöfer et al., 2008; Morales et al., 2011; Rodriguez-Fornells & Münte, 2016; Salamoura & Williams, 2007). Our training study allowed us to examine this question from the earliest stages of acquisition. In addition, our design offers the distinct advantage of presenting only the newly learned language. In several studies that have shown GCE, participants had to actively process their native and second language simultaneously due to task requirements (switching between languages, translating, or ignoring embedded L1 words during L2 production; Bordag & Pechmann, 2007, 2008; Costa et al., 2003; Rodriguez-Fornells & Münte, 2016; Salamoura & Williams, 2007). The necessity to maintain both languages active may have played a role in evoking gender congruency effects in these studies. This cannot be claimed for the present results. Indeed, our study did not require L2 participants to overtly produce or consciously activate their L1. Nonetheless, that the L1 lexicon, and more specifically the grammatical features of L1 candidates, became active during L2 processing was readily apparent in the ERP data. Our results are in line with those reported by Boutonnet et al. (2012) who showed modulation of a late negative component as a function of whether triplets of words, presented in English, all shared the same lexical gender in the participants' native language, Spanish. Thus, as in the current study, even in conditions where the L1 was physically absent, it played a significant role in processing. More specifically, native speakers of gendered languages automatically and irrepressibly activate the L1 gender of nouns, even when processing exclusively in the L2.

In many studies, gender congruency effects have been reported both within language and across languages when participants were required to produce a definite article prior to the noun (Costa et al., 2003; Salamoura & Williams, 2007); although others have shown a GCE when participants produced bare nouns (Bordag & Pechmann, 2007). In light of this, it is of interest to note that while our participants did not produce the lexical items, all auditory nouns were preceded by the definite determiner (e.g. [o parafuso] “the screw” and [a vassoura] “the broom”). It is an empirical question whether the gender congruency effects that we obtained would occur in the absence of the determiner. Given that our participants showed evidence that they had acquired the gender of the L2 nouns, it is possible that they may have retrieved this information during processing, either from a stored representation of the noun or from the morphology of the word form itself, which may then have been the source of interference. Further work is necessary to determine the locus of the interference we found.

Our results show rapid learning of the L2 vocabulary, as demonstrated by both the establishment of an N400 response to newly learned words and by ceiling level accuracy at post-test. Concerning the cortical response, modulations of the N400 have been noted in association with L2 learning and/or artificial languages in several studies. This has been demonstrated in L2 studies that have used longitudinal designs to investigate changes in cortical activity over the course of learning (Chun, Choi, & Kim, 2012; McLaughlin, Osterhout, & Kim, 2004; Stein et al., 2006; Yum, Midgley, Holcomb & Grainger, 2014). Artificial language paradigms have also been used to ascertain whether novel L2 words can be associated quickly with meaning. Breitenstein and colleagues (Breitenstein et al., 2007) used associative learning, where a spoken word was paired with the image of an object with increasing statistical probability over multiple trials. They found that after 5 days of training, newly learned words facilitated the processing of related pictures (in the form of shorter response latencies), indicating integration into existing lexical networks. A similar magnetoencephalography (MEG) study by Dobel et al. (2010) showed a reduction in the mN400 (the MEG component comparable to the ERP N400 component) to correct images preceded by trained spoken words from pre to post-training, indicating that trained words had become associated with existing conceptual representations. Our results corroborate these findings, showing the acquisition of a small L2 vocabulary and understanding of simple sentences following three 30-minute sessions and one 40-minute review session over the course of 4 days, as manifested by the establishment of an N400 response from pre to post-test and increased accuracy, from chance to ceiling level.

Our study demonstrates that learners were sensitive to the grammatical gender of the newly learned words, despite there being no formal or explicit instruction concerning gender and even though the stimuli were only presented aurally. This implies that our participants segmented the auditory stimuli into the determiner and noun and inferred gender information from the properties of the speech signal. Segmenting the audio signal into its syntactic elements is notoriously difficult during second language acquisition (Altenberg, 2005), which is why we created a game that specifically segmented the auditory sentences and required participants to recognize the meanings of the different elements and assemble them in the correct order to recreate the auditory sentence. However, while determiner phrases and verbal phrases were explicitly segmented, the determiner phrases were presented as a single unit (e.g. /okaSimbo/ "the pipe" and /asaia/ "the skirt"). Learners could, in theory, have interpreted the determiner phrase as a whole rather than segment it into the determiner and noun, as indeed there is evidence for in young children. For example, in early stages of acquisition French children may produce forms that reveal segmentation errors (e.g. "le loiseau" and "le noiseau" stemming from the speech signal "l'oiseau" and "un oiseau" "the bird / a bird" Clark, 2009). However, the regularity of the morphological form for the determiner (/a/ or /o/) preceding the noun and the concurring final phoneme of nouns reduces this possibility, as discussed below. Moreover, the majority of participants had formally learned Spanish as a second language throughout secondary school; this may well have prompted them to capitalize on the partial overlap of gender concord rules in Spanish and Portuguese (cf. Brooks & Kempe, 2013).

Various studies with either natural or artificial languages have shown that following implicit training, adult learners rapidly deduce the rules that govern grammatical gender assignment (Öttl & Behne, 2017) and gender concord (Denhovska & Serratrice, 2017; Morgan-Short et al., 2010). Concerning gender assignment, our materials provided a clear phonetic cue, as outlined above, but no semantic information was associated with the gender of nouns. This differs from the materials learned in an artificial language (Öttl & Behne, 2017) in which gender suffixes on the noun were determined by the biological gender of the stimuli. Concerning gender concord, Morgan-Short et al. (2010) found no difference in learning as measured by behavioral (dprime) or cortical sensitivity (P600 response) to determiner-noun gender concord violations as a function of the type of training (implicit or explicit) at the end of training. Using a miniature set of Russian nouns and adjectives, in which adult learners were exposed to noun-adjective gender concord in short written sentences, Denhovska & Serratrice (2017) showed that even under implicit learning, where emphasis was placed on learning the meanings of the sentences and no mention was made of the underlying grammatical rules, participants readily acquired these rules. Moreover, no difference in behavioral accuracy was found for grammatical judgments as a function of the type of instruction (implicit or explicit), although only those who received explicit instruction were able to produce the grammatical rules governing gender concord above chance level. The present results are in line with those found in the above studies, showing that participants rapidly acquire grammatical gender concord rules in a newly learned language following short training sessions, even in the absence of formal instruction.

The pattern of results we obtained suggests that gender congruency effects play an early role during lexical access. This question, i.e. whether grammatical gender exerts an early influence on lexical access or only later, during lexical selection, has been examined in various monolingual studies. Eye-tracking experiments have demonstrated that both children and adults use gender agreement to predict nouns when they are preceded by a gender-marked determiner (Brouwer, Sprenger & Unsworth, 2017; Cholewa, Nitzel, Bürsgens & Günther, 2019; Lew-Williams & Fernald, 2007). Far less evidence of this has been found in the second language (Hopp & Lemmerth, 2018; Lemmerth & Hopp, 2019). In a primed lexical decision task using auditory homophone primes and orthographic targets, Spinelli and Alario (2002) found that gender-marked determiners constrained lexical access to the gender congruent candidate for French homophones. However, subsequent work provided evidence that grammatical gender does not in fact constrain lexical access but acts at a later stage during the selection of the appropriate candidate (Spinelli, Meunier & Seigneuric, 2006). Our results clearly demonstrate that the L1 gender of stored words was activated and interacted with the L2 gender of actually presented words. It remains to be determined whether such was due to the presence of the salient and reliable morphological marking carried in the determiner (i.e. whether participants computed gender based on morpho-phonological cues) or due to the activation of the L2 gender from the auditory noun itself (i.e. retrieval of gender from a newly stored representation). Nevertheless, our results suggest that gender congruency affected lexical access for newly learned L2 nouns, which was hindered when competing gender features from the L1 were activated.

Last, we can note that our design did not allow us to examine the effect of gender concord or gender congruency within the L2. Several studies are warranted to delve further into this question. It would be of interest, for example, to test whether we could establish gender congruency effects within the newly learned L2 by adopting a paradigm similar to that used by Boutonnet et al. (2012), in which successive trials carried the same L2 gender, or a visual world paradigm (Hopp & Lemmerth, 2018) in which the gender of the items is varied both across languages and within the L2. Concerning gender concord, a typical violation paradigm could be added to the current design, whereby the learned L2 nouns are preceded by either the correct determiner or the incorrect determiner (Foucart & Frenck-Mestre, 2011, 2012; Morgan-Short et al., 2010). In this light, a further development of the present work would be to present the learned L2 nouns as bare nouns, in order to determine whether the same pattern of gender congruency interference is observed.

To conclude, we have provided clear electrophysiological evidence of gender congruency effects in an L2, from the very beginning of acquisition. The N400 effect elicited by mismatched compared to matched audio-visual pairs for gender congruent trials was basically annulled for gender incongruent trials. To our knowledge, no prior work has provided evidence of GCE, either behaviorally or via the cortical response to newly learned L2 words. It is important to note that our behavioral results clearly demonstrate that participants learned the correct association between auditory words and images and that this was independent of both the gender in the second language vocabulary and, crucially, the congruency of gender across the L2 and the participants' L1. Hence, we have also provided evidence of the value of a multi-disciplinary approach to bring the effects of cross-linguistic gender congruency to light, at least in the early stages of acquisition.

References

Alario, F.-X., & Ferrand, L. (1999). A set of 400 pictures standardized for French: Norms for name agreement, image agreement, familiarity, visual complexity, image variability, and age of acquisition. *Behavior Research Methods, Instruments, & Computers*, *31*(3), 531–552.

Altenberg, E.P. (2005). The perception of word boundaries in a second language. *Second Language Research*, 21, 325-358.

Arnon, I., & Ramscar, M. (2012). Granularity and the acquisition of grammatical gender: How order-of-acquisition affects what gets learned. *Cognition*, *122*(3), 292–305.

Barber, H., & Carreiras, M. (2005). Grammatical Gender and Number Agreement in Spanish: An ERP Comparison. Journal of Cognitive Neuroscience, 17, 137-155.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Bigi, B. (2015). SPPAS - Multi-lingual approaches to the automatic annotation of speech. *The Pho*netician, *111– 112*(ISSN:0741-6164), 54–69.

Boersma, P. & Weenink, D. (2018). Praat: doing phonetics by computer [Computer program]. Version 6.0.37, retrieved 14 March 2018 from http://www.praat.org/

Bordag, D. (2004). Interaction of L1 and L2 systems at the level of grammatical encoding: Evidence from picture naming. *EUROSLA Yearbook*, *4*(1), 203–230.

Bordag, D., & Pechmann, T. (2007). Factors influencing L2 gender processing. *Bilingualism: Language and Cognition, 10*, 299–314.

Bordag, D., & Pechmann, T. (2008). Grammatical gender in translation. *Second Language Research, 24*, 139–166.

Boutonnet, B., Athanasopoulos, P., & Thierry, G. (2012). Unconscious effects of grammatical gender during object categorisation. *Brain Research*, *1479*, 72–79

Breitenstein, C., Zwitserlood, P., de Vries, M., Feldhues, C., Knecht, S., & Dobel, C. (2007). Five days versus a lifetime: Intense associative vocabulary training generates lexically integrated words. *Restorative Neurology and Neuroscience*, *25*(5,6), 493–500.

Brooks, P. J., & Kempe, V. (2013). Individual differences in adult foreign language learning: The mediating effect of metalinguistic awareness. *Memory & Cognition*, 41(2), 281–296.

Brouwer, S., Sprenger, S., and Unsworth, S. (2017). Processing grammatical gender in Dutch: evidence from eye movements *Journal of Experimental Child Psychology*, 159, 50–65.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, *14*(1), 177-208.

Cholewa, J., Neitzle, Bürsgens & Günther, T. (2019). Online-processing of grammatical gender in noun-phrase decoding: An eye-tracking study with monolingual German 3rd and 4th graders. Frontiers in Psychology, 10, Article 2586.

Chun, E., Choi, S., & Kim, J. (2012). The effect of extensive reading and paired-associate learning on long-term vocabulary retention: An event-related potential study. *Neuroscience Letters*, 521(2), 125–129.

Clark, E.V. (2009). First Language Acquisition. 2nd Edition. Cambridge University Press.

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(5), 1283–1296.

Costa, A., Kovacic, D., Franck, J., & Caramazza, A. (2003). On the autonomy of the grammatical gender systems of the two languages of a bilingual. *Bilingualism: Language and Cognition*, *6*(3), 181–200.

Dahan, D., Swingley, D., Tanenhaus, M., & Magnuson, J. (2000). Linguistic Gender and Spoken-Word Recognition in French. *Journal of Memory and Language - J MEM LANG*, *42*, 465–480.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283–321.

Denhovska, N., & Serratrice, L. (2017). Incidental Learning of Gender Agreement in L2. *Journal of Psycholinguistic Research*, *46*(5), 1187–1211.

Dobel, C., Junghöfer, M., Breitenstein, C., Klauke, B., Knecht, S., Pantev, C., & Zwitserlood, P. (2010). New names for known things: On the association of novel word forms with existing semantic information. *Journal of Cognitive Neuroscience*, *22*(6), 1251–1261.

Dowens, M. G., Guo, T., Guo, J., Barber, H., & Carreiras, M. (2011). Gender and number processing in Chinese learners of Spanish—Evidence from Event Related Potentials. *Neuropsychologia*, *49*(7), 1651–1659.

Foucart, A., & Frenck–Mestre, C. (2012). Can late learners acquire new grammatical features? Evidence from ERPs and eye-tracking. *Journal of Memory and Language*, 66, 226-248.

Foucart, A., & Frenck–Mestre, C. (2011). Grammatical gender processing in L2: Electrophysiological evidence of the effect of L1 -L2 syntactic similarity. *Bilingualism, Language and Cognition*, 14, 379-399.

Frenck–Mestre, C. (2005). Eye-movement recording as a tool for studying syntactic processing in a second language: a review of methodologies and experimental findings. *Second Language Research*, 21, 175-198.

Gunter, T., Friederici, A., & Schriefers, H. (2000). Syntactic Gender and Semantic Expectancy: ERPs Reveal Early Autonomy and Late Interaction. *Journal of Cognitive Neuroscience*, 12, 556-568.

Hagoort, P. (2003). Interplay between Syntax and Semantics during Sentence Comprehension: ERP Effects of Combining Syntactic and Semantic Violations. *Journal of Cognitive Neuroscience*, 15(6), 883–899.

Hopp, H. & Lemmerth, N. (2018). Lexical and syntactic congruency in L2 predictive gender processing. *Studies in Second Language Acquisition*, 40, 171-199.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology*, *62*, 621–647.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(1), 1–26.

Lemhöfer, K., Spalek, K., & Schriefers, H. (2008). Cross-language effects of grammatical gender in bilingual word recognition and production. *Journal of Memory and Language*, *59*(3), 312–330.

Lemmerth, N., & Hopp, H. (2019). Gender processing in simultaneous and successive bilingual children: Cross linguistic lexical and syntactic influences. *Language Acquisition*, *26*(1), 21–45.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1–38.

Lew-Williams, C., & Fernald, A. (2007). Young Children Learning Spanish Make Rapid Use of Grammatical Gender in Spoken Word Recognition. *Psychological Science*, *18*(3), 193–198.

McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: Minimal instruction produces rapid change. *Nature Neuroscience*, *7*(7), 703–704.

Molinaro, N., Barber, H. A., Caffarra, S., & Carreiras, M. (2015). On the left anterior negativity (LAN): The case of morphosyntactic agreement: a reply to Tanner et al. *Cortex*, *66*, 156–159.

Morales, L., Paolieri, D., & Bajo, T. (2011). Grammatical Gender Inhibition in Bilinguals. *Frontiers in Psychology*, *2*.

Morales, L., Paolieri, D., Dussias, P. E., Valdés Kroff, J. R., Gerfen, C., & Bajo, M. T. (2016). The gender congruency effect during bilingual spoken-word recognition. *Bilingualism*, *19*(2), 294–310.

Morgan-Short, K., Sanz, C., Steinhauer, K., & Ullman, M. T. (2010). Second Language Acquisition of Gender Agreement in Explicit and Implicit Training Conditions: An Event-Related Potential Study. *Language Learning*, *60*(1), 154–193.

New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 516–524.

Nichols, E. S., & Joanisse, M. F. (2016). Functional activity and white matter microstructure reveal the independent effects of age of acquisition and proficiency on second-language learning. *NeuroImage*, *143*, 15–25.

Nickels, L., Biedermann, B., Fieder, N., & Schiller, N. (2015). The lexical-syntactic representation of number. *Language, Cognition and Neuroscience*, *30*, 287–304.

Öttl, A., & Behne, D. M. (2017). Assessing the Formation of Experience-Based Gender Expectations in an Implicit Learning Scenario. *Frontiers in Psychology*, *8*.

Popov, S. & Bastiaanse, R. (2018). Processes underpinning gender and number disagreement in Dutch: An ERP study. *Journal of Neurolinguistics*, 46, 109-121.

Popov, S., Miceli, G., Ćurčić-Blake, B., & Bastiaanse, R. (2020). The role of semantics and repair processes in article noun gender disagreement in Italian: An ERP study. *Brain and Language*, *206*, 104787.

Rodriguez-Fornells, A., & Münte, T.F. (2016) Syntactic interference in bilingual naming during language switching: An electrophysiological study. In J.W. Schweiter (Ed.) Cognitive Control and Consequences of Multilingualism, John Benjamins, Amsterdam/Philidelphia. pp. 239-270.

Sá-Leite, A. R., Fraga, I., & Comesaña, M. (2019). Grammatical gender processing in bilinguals: An analytic review. *Psychonomic Bulletin & Review, 26*(4), 1148–1173.

Sabourin, L., & Stowe, L. A. (2008). Second language processing: When are L1 and L2 processed similarly. *Second Language Research*, 24, 397–430.

Sabourin, L., Stowe, L.A., & de Haan, G.J. (2006). Transfer effects in learning a second language grammatical gender system. *Second Language Research*, 22, 1–29

Salamoura, A., & Williams, J. N. (2007). The representation of grammatical gender in the bilingual lexicon: Evidence from Greek and German. *Bilingualism: Language and Cognition*, *10*(3), 257–275.

Schiller, N., & Caramazza, A. (2006). Grammatical gender selection and the representation of morphemes: The production of Dutch diminutives. *Language and Cognitive Processes*, *21*, 945–973.

Snodgrass, J. G., & 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), 174–215.

Spinelli, E., & Alario, F.-X. (2002). Gender Context Effects on Homophone Words. *Language and Cognitive Processes*, 457–469.

Spinelli, E., Meunier, F., & Seigneuric, A. (2006). Spoken word recognition with gender-marked context. *The Mental Lexicon*, 277–297.

Stein, M, Dierks, T., Brandeis, D., Wirth, M., Strik, W., & Koenig, T. (2006). Plasticity in the adult language system: A longitudinal electrophysiological study on second language learning. *NeuroImage*, *33*(2), 774–783.

Tanner, D. (2015). On the left anterior negativity (LAN) in electrophysiological studies of morphosyntactic agreement: A Commentary on “Grammatical agreement processing in reading: ERP findings and future directions” by Molinaro et al., 2014. *Cortex*, *66*, 149–155.

Tokowicz, N., & MacWhinney, B. (2005). Implicit and Explicit Measures of Sensitivity to Violations in Second Language Grammar: An Event-Related Potential Investigation. *Studies in Second Language Acquisition*, *27*(2), 173–204.

Wicha, N., Moreno, E., & Kutas, M. (2004). Anticipating Words and Their Gender: An Event-related Brain Potential Study of Semantic Integration, Gender Expectancy, and Gender Agreement in Spanish Sentence Reading. *Journal of Cognitive Neuroscience*, *16*, 1272–1288.

Yum, Y., Midgley, K., Holcomb, P., & Grainger, J. (2014). An ERP study on initial second language vocabulary learning. *Psychophysiology*, 51(4), 364–373.

Chapter 3

What role does motor activation play in action language processing? An EEG study

Zappa, A., Bolger, D., Pergandi, J.-M., Ghio, A., & Frenck-Mestre, C., *under revision*

Neuroimaging and behavioral evidence points to the recruitment of sensory-motor systems during semantic access. However, much debate remains surrounding the role of motor activations during language processing. In the current study we manipulated motor and semantic compatibility while measuring participants' cortical activity using electroencephalography (EEG). Participants listened to action sentences indicating a movement away from or towards one's body and accepted them by performing a compatible or incompatible action. We measured motor-related cortical activity through time-frequency and motor and language-related ERPs (event-related potentials) during the auditory processing of action sentences. ERP analyses showed a greater negative deflection of the N400 for compatible versus incompatible trials, suggesting an inhibitory effect of compatible motor processes on action verb comprehension. Results also provide evidence of action-related mu and beta suppression at centro-parietal sites during action sentence processing. This study adds to an important new vein in cognition research that, rather than focusing on the embodiment vs disembodied debate, prioritizes determining the exact role of motor activation in cognition. Although previous behavioral studies have shown interference when compatible motor and semantic processes overlap, ours is the first EEG study to provide cortical measures of this effect.

3.1 Introduction

How independent are linguistic representations? Several recent studies have suggested that the representations of lexical items are distributed and incorporate motor as well as linguistic features. Studies involving both single word processing (Pulvermüller 1999, 2002, 2012; Pulvermüller & Fadiga, 2010) and sentence processing (van Elk, van Schie, Zwaan & Bekkering, 2010) have revealed that semantic representations of words activate sensory-motor areas in addition to classic language regions. However, the role of these activations is complex and requires further investigation. Herein, we discuss various studies that have sought to better understand the precise role of motor activations in language processing by examining interactions between action and action-related language processing.

When considering the possibility that semantics is grounded in sensory-motor systems, it is important to keep in mind that meaning as conveyed by language is highly contextual. Indeed, several studies have revealed motor-semantic interactions to be context-dependent. For instance, they are heavily influenced by linguistic polarity, or whether processed sentences are affirmative or negative (Alemanno et al., 2012; Aravena et al., 2012), and linguistic focus, or whether the focus of the sentence is on the action being performed or on the mental state of the performer (Aravena et al., 2014). Furthermore, the level of complexity of movement sequences impacts the direction of the effect (Shebani & Pulvermüller, 2018) as do task demands, especially as concerns whether a semantic decision is required or not (Papeo, Vallesi, Isaja & Rumiati, 2009; Sato, Mengarelli, Riggio, Gallese & Buccino, 2008). A crucial aspect of semantic-to-motor and motor-to-semantic effects that emerge from the combination of motor response and action-related language is the temporal relationship between the two. The literature has in fact shown that the precise moment when action language is processed in relation to movement onset can determine whether semantic-to-motor effects are inhibitory, facilitatory or absent (Boulenger et al., 2006; Sato et al., 2008; for a review see García & Ibáñez, 2016). The present study aimed to elucidate the complex role of the motor system in representing and retrieving meaning by using behavioral measures and electroencephalography (EEG) to examine the overlap of motor response and processing the meaning of action language.

Evidence that linguistic representations involve sensory-motor systems largely stems from neuroimaging studies of single word processing. These studies show a somato-specific overlap between motor processes and the semantic processing of action-related verbs such that motor circuits that subserve executing or observing actions involving the mouth, hands or legs are also implicated in processing verbs describing those actions (Aziz-Zadeh, Wilson, Rizzolatti & Iacoboni, 2006; Hauk, Johnsrude & Pulvermüller, 2004; Tettamanti et al., 2005; see Shebani & Pulvermüller, 2018 for a review). In a recent fMRI (functional magnetic resonance imaging) experiment, Horoufchin, Bzdok, Buccino, Borghi & Binkofski (2018) found that observing actions compared to static objects produced different neural activation patterns, as did reading verbs compared to nouns. They concluded that visual recognition and language comprehension involve simulation processes that use canonical neurons (activated when graspable objects and object-directed actions are seen) to distinguish objects from actions and mirror neurons (activated when a goal-specific action is performed or observed) to distinguish nouns from verbs.

Behavioral studies have also shown interactions between motor and linguistic systems via action and action language compatibility paradigms (Aravena et al., 2010; Gianelli, Farnè, Salemme, Jeannerod & Roy, 2011; Glenberg & Kaschak, 2002; Glenberg et al., 2008). A widely employed paradigm is the Action-sentence compatibility effect (ACE) paradigm, which requires participants to make a sensibility judgment to a sentence describing an action. At play is the overlap between the behavioral response and the semantic representation. Indeed, the physical action used to respond is either compatible or incompatible with the action described in the sentence. This task was first developed by Glenberg & Kaschak (2002). In their experiment, participants read sensible and nonsense sentences and judged their sensibility as quickly as possible. Sensible sentences all implied an action away from (e.g. “You gave the pizza to Andy”) or towards (e.g.“Andy gave you the pizza”) the body. There were two answering conditions: yes-is-far, for which participants had to extend their arm to answer “yes”, and yes-is-near, for which they had to move their arm toward their body to answer “yes”. Results showed an interaction between the direction described in the sentence and the direction of the arm movement performed such that reading times were faster and arm movements were facilitated when the answering direction was compatible with the direction described in sentences. For example, when processing the sentence “You gave the pizza to Andy” participants answering in the “yes-is-far” condition, considered compatible with the sentence, had shorter reading times and faster responses than when they read “Andy gave you the pizza”. Results were interpreted as showing that participants performed a mental simulation of the sentence to comprehend it. Following this seminal experiment, variations of the ACE paradigm have shown the effect using different hand and arm movements (Diefenbach, Rieger, Massen & Prinz, 2013; Dudschig, de la Vega & Kaup, 2014; Zwaan & Taylor, 2006).

Despite the growing number of neurobiological and behavioral studies that reveal motor-language interactions, the interpretation that these effects indicate a functional relationship between motor activation and language processing is by no means unanimous (Mahon

& Caramazza, 2008). Indeed, in the last ten years, studies have shown that motor-semantic interactions are largely dependent on linguistic context. For instance, action verbs produced motor resonance in sentential contexts only when embedded in affirmative sentences; negative sentences with the same verbal content produced no reliable variation in response (Aravena et al., 2012; but see Alemanno et al., 2012). This suggests that motor activation is fact not *necessary* for semantic access and may only occur post lexical access. This hypothesis was bolstered by a subsequent study showing that motor activations were only observed for sentences in which the action verb was the focus ("John signs the contract"), as opposed to sentences that focused on the agent's mental state ("John wants to sign the contract") (Aravena et al., 2014). In similar fashion, when verbs were presented in isolation ("kick") or in literal sentences ("kick the ball") the motor cortex was activated, but not when they were embedded in idioms ("kick the bucket") (Raposo, Moss, Stamatakis & Tyler, 2009; but see Boulenger, Hauk & Pulvermüller, 2009 for opposite results). Finally, linguistic perspective has also been shown to constrain the ACE. Gianelli and colleagues (2011) found an ACE for sentences in the first person ("You gave the pizza to Louis"), but not in the third person ("Lea gave a pizza to Louis"). Nonetheless, the ACE was restored for third person sentences when a virtual body was added to the stimulation screen, allowing participants to position themselves when taking the third person perspective.

Importantly, several behavioral studies have explored the temporal dynamics of the ACE and shown that processing action-related language can either interfere with or facilitate the execution of motor tasks (de Vega, Moreno & Castillo, 2013; Boulenger et al., 2006; Diefenbach et al., 2013; Kaschak & Borreggine, 2008; Sato et al., 2008; for a review see García & Ibáñez, 2016). The direction of the effect is contingent on whether linguistic processing overlaps with or precedes the motor task. Specifically, compatibility between action-related language and answer movement hinders responses (negative ACE) during early windows, from 0-400 msec for single verbs and up to 1 sec when sentence comprehension is required. However, if response planning occurs at sentence onset or if movement onset is delayed to after the action sentence is fully processed, the manual response is facilitated (positive ACE). García and Ibáñez (2016) stipulate that when motor and linguistic processes occur simultaneously, interference arises due to competition for shared neural resources. In contrast, when movement is planned outside of this overlapping window, it leads to priming effects that likewise point to shared neural resources.

Evidence that linguistic context and linguistic perspective dramatically affect whether or not the motor system is activated concurrently with language comprehension indicates that the ACE is more constrained than was originally thought. Moreover, the fact that the direction of the ACE depends on timing highlights the need to test specific aspects of the effect, notably that of functionality. Recently, Papesh (2015) performed a series of experiments that originally aimed to 1) determine whether the ACE could be found using a novel mouse-tracking method and 2) examine whether higher-level cognitive processes such as disembodiment would influence the ACE. In three novel designs she failed to extend the ACE to sentential contexts. Moreover, in five subsequent experiments that directly replicated previous studies, using the same stimuli and highly similar response methods as the seminal work (Glenberg & Kaschak, 2002), no ACE effects were found. Following a Bayesian analysis of the ensemble of previous studies, Papesh (2015) concluded that the evidence for behavioral ACE effects is weak at best.

Focusing on when motor resonance occurs in the brain in relation to semantic processing can help reveal whether it has functional significance for linguistic processing. An effective tool to examine this question is EEG, given its high temporal resolution and capacity to chart cortical activation. A component that is often used in language studies is the N400, a negative-going wave usually observed between 300-500 msec after stimulus onset at centro-posterior (but sometimes also fronto-central) electrodes over both hemispheres (Kutas & Federmeier, 2000; Kutas & Hilliyard, 1980). It has generally been linked to semantic processes such as lexical access and lexical-semantic integration (Kutas & Hilliyard, 1980; Kutas & Federmeier, 2011). Also relevant to studies investigating embodied effects, two motor potentials thought to reflect motor cortex activity are the MP, a motor potential starting around -90 msec before response onset, and the RAP, a re-afferent potential occurring 200-300 msec following movement onset (Aravena et al., 2010).

Aravena and colleagues (2010) recorded event-related potentials (ERPs) in an ACE paradigm to examine the precise timing of the bidirectional impact between motor processes and language comprehension, focusing on motor components (MP and RAP) and the N400. Participants heard sentences containing a final action verb and answered with a hand shape that either matched the hand shape described by the verb or constituted a mismatch. Motor preparation and execution through movement-related cortical potentials were measured at specific electrodes. Results showed an enhanced amplitude for both the motor potential (MP) (-90 msec before response onset) and the re-afferent potential (RAP) (200–300 msec after response onset) for compatible versus incompatible trials. The early MP response suggests that action sentence compatibility facilitated movement preparation due to semantic priming. Moreover, incompatible trials produced an N400-like effect compared to compatible trials, although effects were limited as concerns both timing and distribution. The fact that motor preparation was bolstered by meaning-action congruity, whereas semantic integration was hampered by the meaning-action incongruity, was interpreted as illustrating a robust ACE, thereby indicating that meaning influenced motor programming prior to lexical access and that motor and semantic processes likely belong to a common network.

Subsequent ERP studies have examined this issue by manipulating sensory-motor information or affordances in sentential contexts and focusing on the N400 component. N400 effects in these studies have been presented as evidence supporting an embodied approach to language processing whereby sensory-motor information influences semantic processes. For instance, the cost of switching sensory modalities during sentence comprehension ("The cellar is dark" (visual) vs "A mitten is soft" (tactile)) has been shown to correlate with increased N400 amplitudes (Collins, Pecher, Zeelenberg & Coulsen, 2011; Hald, Marshall, Janssen & Garnham, 2011). Similarly, studies that manipulated the affordances of the noun in relation to the verb (e.g. "The boys searched for bushes/branches with which they went drumming") found that these physical/semantic affordances produced an N400 effect (Chwilla, Kolk & Vissers, 2007).

ERPs are nonetheless limited when it comes to measuring motor activation as they use phase-locked linear averaging methods, which do not capture the part of the motor cortex signal that is not phase-locked to the stimulus (Pfurtscheller & Lopes da Silva, 1999; Vukovic & Shtyrov, 2014). In contrast, time-frequency analysis of the EEG signal takes into account both phase-locked and non phase-locked cortical activity. Hence, motor activity during language processing can be measured through a pattern of event-related desynchronization (ERD), or power suppression, and synchronization (ERS), or power enhancement, in the time interval following stimulus onset compared to a pre-stimulus baseline. Event-related spectral perturbation (ERSP) (Grandchamp & Delorme, 2011; Makeig, 1993; Makeig, Debener, Onton & Delorme, 2004) groups ERD and ERS and calculates the power spectrum on a trial-by-trial basis. This method is widely used to study the cortical response to motor events (for a review see Hobson & Bishop, 2016). Specifically, the suppression of mu (8–13 Hz), and beta (13–30 Hz) rhythms, synchronized patterns of electrical activity recorded over the sensory-motor cortex, is associated with performing and observing movement (Caetano, Jousmäki & Hari, 2007; Koelewijn, van Schie, Bekkering, Oostenveld, & Jensen, 2008; Pfurtscheller & Lopes da Silva, 1999; Pineda, 2005) as well as motor imagery (Matsumoto et al., 2010).

Mu and beta desynchronization have likewise been interpreted as a correlate of facilitated action language processing (Klepp, van Dijk, Niccolai, Schnitzler & Biermann-Ruben, 2019). During a passive reading task, beta desynchronization was observed between 150 and 500 msec after single visual word onset for hand and foot action verbs, in comparison to non-action verbs, as well as later-stage alpha-band desynchronization for hand versus non-action verbs (Niccolai et al., 2014). Zappa and colleagues went one step farther and used a Go-Nogo task to examine the influence of planning real actions on auditory action verb processing using time-frequency measures (Zappa et al., 2019). During Go trials, participants performed real actions on virtual objects in a realistic 3D virtual environment. Go trials elicited greater mu and beta desynchronization in the time window associated with semantic processing compared to Nogo trials, due to inhibited motor preparation during the processing of action verbs. Beyond single-verb processing, action-related sentences have also been shown to produce mu and beta desynchronization, or motor resonance (van Elk et al., 2010). Unlike the ACE study described earlier, which showed no motor activation for negative sentences (Aravena et al., 2012), negative hand-action related sentences produced greater mu ERD than their positive counterparts (Alemanno et al., 2012). Another study showed greater mu and low beta (15-20 Hz) desynchronization for action-related (e.g. “You will cut the strawberry cake”) sentences compared to perceptive (e.g. “You will notice the bright day”) and abstract sentences (e.g. “You will doubt the patient’s argument”) (Moreno et al., 2015). However, maximal mu desynchronization was not found at the verb itself but later in the sentence, at the first and second noun. These results suggest that motor activation occurred as a result of sentence integration and was not simply related to the lexical processing of the action verb, in line with the idea that action language comprehension involves simulation (Boulenger et al., 2009). Importantly, they point to the effectiveness of mu and beta ERD as good indicators of motor activation, including during language processing, thereby legitimizing their use in studies of embodied semantics.

In the present study we built on the relatively scarce body of EEG work that manipulated action sentence meaning and movement compatibility to examine the influence of motor processes on semantic representation. Notably, we used both event-related potentials and time-frequency to measure the effect of motor-semantic interactions during the processing of action-related language, using an ACE paradigm. Despite the substantial number of behavioral ACE studies (Boulenger et al., 2006; de Vega & Urrutia, 2011; Diefenbach et al., 2013; Glenberg & Kaschak, 2002; Kaschak & Borreggine, 2008; Sato et al., 2008) and the growing number of studies using time-frequency to measure motor resonance during sentence processing (Lam, Bastiaansen, Dijkstra & Rueschemeyer, 2017; Moreno et al., 2015), very few studies have combined these two measures to investigate the relationship between cortical motor resonance and action sentence processing (de Vega et al., 2016).

We recorded EEG in a behavioral ACE paradigm that required participants to perform arm movements at the end of auditory sentences describing actions that were concordant with or opposite to the movement. As in Glenberg and Kaschak’s (2002) study, participants performed a semantic task whereby they accepted sentences by either extending their arm away from their body or bringing it toward themselves. We manipulated semantic-motor compatibility such that the action described in the sentences (e.g.“Suzanne a pris les cartes et les a distribuées” [Suzanne took the cards and distributed them]) was either compatible with the performed action (in this case, extending one’s arm away from one’s body) or not (bringing one’s arm towards one’s body). Importantly, acceptability was only defined by the sentence-final verb and hence motor planning could only occur at the end of the sentence, simultaneously to action language processing. Unlike Aravena et al.’s study (2010), where a repetitive movement was performed, in our study participants were required to choose between a compatible and an incompatible action, building on Aravena and colleagues' methodology in order to better control the overlap between motor preparation and action language processing. We used third-person sentences to avoid overt imagery processes, which could in themselves account for motor resonance during language comprehension (Aravena et al., 2010). To investigate motor-semantic interactions, we focused on movement-related potentials (MP and RAP) and later components linked to semantic integration (N400). We observed direct motor resonance during language processing through time-frequency analysis on EEG oscillations in the mu (8–13 Hz) and upper (20–30 Hz) and lower (14-20 Hz) beta bands.

In sum, the current study aimed at developing the literature on the brain markers of the the ACE by further approximating motor and semantic processes compared to previous studies (Aravena et al., 2010). In line with behavioral studies showing that the direction of the ACE is contingent on when motor planning occurs in relation to action-language processing (for a review see García & Ibáñez, 2016), we predicted that, given the overlap between motor preparation and action verb processing, if motor and semantic processes share common mechanisms, compatible trials would show interference in the form of an increased N400 amplitude or decreased motor preparation (MP and RAP) and motor resonance (mu and beta desynchronization).

3.2 Methods

3.2.1 Participants

Thirty right-handed French native speakers (aged 19–27, sd = 2.82, 19 females) participated in the study. Four participants were excluded from analyses due to poor signal quality, leaving 26 participants in the final sample (aged 19-27, sd=2.95, 16 females). Participants were volunteers from the student population of the Aix-Marseille Université. They had no history of neurological insult and received monetary compensation for their participation. They were debriefed about the purpose of the experiment at its end. All participants gave their written informed consent prior to the experiment, and the study was approved by the local university ethics committee.

3.2.2 Materials

Experimental stimuli consisted of 40 auditory sentences in French (see Annex). Twenty verbs denoting an action that involved an arm movement either toward or away from the body

(10 for each action) were used to create 20 "toward" (TW) ("Emilie a pris son verre de vin et l'a bu" [Emilie picked up her glass of wine and drank it]) and 20 "away" (AW) sentences ("Suzanne a pris les cartes et les a distribuées" [Suzanne took the cards and distributed them]). Each verb was used twice, in independent sentences. To validate the stimuli, we conducted an online pre-test including 20 participants who did not take part in the main study. They were asked to evaluate the movement implied by each sentence on a 5-point Likert scale (1 denoting completely towards the body and 5 completely away from the body). Only sentences that received a score of 5 for 90% of the participants for the "away" sentences and, in complementary fashion, of 1 for "toward" sentences were retained for the purposes of the experiment. All sentences were in the passé composé tense, which generally corresponds to the simple past in English, and contained a two main clauses, with the critical verb in the second one. The critical verbs were presented in past participle form preceded by a clitic object and were always sentence-final ("Marie a pris la pièce et l'a jetée" [Mary took the coin and tossed it]). The pronominal referent, presented in the main clause, was always human (half male). All experimental action verb sentences were presented in both compatible and incompatible conditions as defined by the direction of movement participants performed to respond (see below). Participants responded positively in only one direction such that in each group, 20 of the 40 test trials were compatible and the other 20 were incompatible with the action performed to accept the sentences.

Sentences were produced by a male native French-speaker and digitally recorded at 48 kHz (32-bit float) in a professional sound booth in a single session. Individual auditory sentences were spliced using Audacity 2.2.1 software. Sentence duration varied from 2000 to 2900 msec (for AW, M=2430, SD=290; for TW, M=2310, SD=190). Each sentence was analyzed using SPASS (Bigi, 2015) to determine the onset of each auditory word. Non-audible triggers were placed at the onset of each auditory word, with the clitic object and auxiliary ("l'a" it + past participle) treated as one element followed by a trigger at the onset of the critical past participle region ("jetée" tossed). Twenty filler sentences were created using the same 20 verbs as those used for the experimental sentences. All fillers were created to be semantically anomalous ("Jules a pris ses neurones et les a distribués" [Jules took his neurons and distributed them]), as verified by an independent sample of 15 participants who rated the semantic acceptability of these sentences in an online questionnaire on a 5 point Likert scale (1 denoted completely unacceptable and 5 completely acceptable). Only sentences that received a score of 1 by 90% of the participants were retained. Four additional sentences (2 sensible and 2 anomalous) were created for the training session. For all participants, a single list of 60 auditory sentences (40 experimental and 20 fillers) was presented in one of four random orders, preceded by the 4 warm-up trials.

3.2.3 Procedure

Participants sat comfortably at a desk situated 60 cm away from a CRT screen in an electrically shielded sound attenuated booth. A trial began with an ocular fixation cross displayed in the center of the computer monitor for 200 msec prior to and for the duration of the auditory sentence, which was presented via electrically shielded speakers. Auditory stimuli were followed by a visual question mark that remained on until the participant's manual response. A visual "blink" prompt was displayed immediately thereafter for 2 seconds. The experimental session lasted roughly 30 minutes, including two breaks. Participants were instructed to judge the semantic acceptability of the sentences via a rectangular response box (33cm. long) equipped with three vertically placed, equally spaced, buttons (4 cm. circumference) labeled "yes", "no" and "next", with the latter in the middle (neutral) position. Participants placed their hand on the response box in the neutral ("next") position at the onset of the experiment. They were instructed to respond using the palm of their right hand then return to the neutral "next" position and press it to initiate the next trial. All experimental sentences were semantically acceptable and hence required a positive response. The "yes" button was located at the top of the response box for half of the participants and at the bottom for the other half, in a way that a positive response required movement either away from or toward their body. At no point were participants informed about the relationship between the direction of hand movements and that implied by the sentences.

3.2.4 ERP Data Acquisition and Processing

EEG activity was recorded continuously from 64 scalp electrodes located at left and right hemisphere positions over frontal, central, parietal, occipital, and temporal areas by means of a 64-channel electrode cap mounted with silver-chloride active electrodes (BioSemi Active Two system AD box). Individual electrodes were adjusted to a stable offset lower than 20mV. EEG data were sampled online at 512 Hz. Blinks and vertical eye movements were monitored via an electrode placed under the right eye and horizontal eye movements were monitored via an electrode placed at the outer canthus of the left eye. One electrode was placed over each mastoid. All electrodes were re-referenced to the average of the two mastoids offline. EEG was recorded continuously during the experiment and periods spanning from -100 msec pre-stimulus onset to 1100 msec post-stimulus onset were used post-recording for analyses. A low pass digital filter of 30 Hz was applied post-recording. Trials contaminated by ocular motor or muscular artifacts were excluded using automated routines that were manually checked. The percentage of trails retained for analyses was 88% for the Compatible condition and 87% for the Incompatible condition, and 85% for the "away" and 83% for the "toward" sentences.

The ERP data were modeled independently in linear mixed effect models for the mean voltage amplitudes in three time windows, in line with prior research (Aravena et al., 2010, 2012). Two components were related to motor mechanisms, namely the MP, which peaks roughly between -90 and 50 msec following movement onset, and the RAP, which peaks around 200-300 msec after movement onset. Both components were time-locked to the offset of the critical verb. In addition, the N400 component was calculated 300-600 msec time-locked to the onset and offset of the critical verb. Analyses for the MP and RAP components were conducted on the data acquired at midline electrodes (Fz, FCz, Cz, CPz).

Analyses for the N400 component were conducted on the data acquired at 35 electrodes, including 5 over midline (Fz, FCz Cz, CPz, Pz), and 24 lateral electrodes divided equally over the left (F1, F3, F5, FC1, FC3, FC5, C1, C3, C5, CP1, CP3, CP5, P1, P3, P5) and right (F2, F4, F6, FC2, FC4, FC6, C2, C4, C6, CP2, CP4, CP6, P2, P4, P6) hemispheres. Electrodes were selected to include fronto-central electrodes based on studies that have investigated motor-semantic effects (Aravena et al., 2010; van Elk et al., 2010).

3.2.5 Event-related Spectral Perturbation (ERSP)

The event-related spectral perturbation (ERSP) was calculated on both the data locked to the verb onset and verb offset, using the FieldTrip toolbox (Oostenveld et al., 2011). To compute the ERSP, time-frequency decomposition was effectuated at the single trial level for each participant and each condition (Compatible and Incompatible) by applying complex Morlet wavelets of 7 cycles over the 4Hz to 35Hz frequency band. This yielded a spectral bandwidth of 1.4Hz and 10Hz at the lowest (4Hz) and highest (35Hz) frequency, respectively. Each trial had a total duration of 3 seconds (-1100 msec to +1900 msec). The trial-level power was averaged to yield a grand-average time-frequency map for each participant. The grand-average post-stimulus power of both the data locked to verb onset and offset was z-score normalized relative to the verb onset pre-trial interval (-200 msec to 0 msec prior to verb onset).

3.3 Results

3.3.1 Behavioral measures

End of sentence responses and associated response times (RTs) were recorded for each participant. We used generalized mixed effects models (glmer) (Kuznetsova, Brockhoff & Christensen, 2017) implemented in R (R Core Team, 2017) to analyze accuracy data for end of sentence responses, including the fixed factors Group (Away vs Toward), Condition (Compatible vs Incompatible) and their interaction, with random intercepts for Participant and Item. For response times, outliers, defined by RTs outside +2 SD per condition were excluded. We applied a linear mixed effects model (lmer) that included the fixed factors Group (Away vs Toward), Condition (Compatible vs Incompatible) and their interaction. Participant and Item both included random intercepts with a random slope of Condition for Participant. The factors were treatment-coded, and the reference levels were Away and Compatible. Models were implemented under R packages lme4 (Bates, Mächler, Bolker, Walker, 2015). Below we report the results from the maximal models (Barr, Levy, Scheepers & Tily, 2013) provided that the models converge.

Accuracy

Accuracy was at ceiling level. No significant variation in accuracy was found as a function of any factor.

Response time

The model revealed an interaction of Condition x Group (β= -248.93, se = 93.18, t =-2.672, p<0.01). The data were subsequently modeled independently for each group. In the Away group the effect of Condition was not significant (β=107.88, se = 66.06, t = 1.633, p = 0.11). In the Toward group, there was a trend for the effect of Condition (β= -142.30, se = 70.53, t = -2.018, p = 0.056) with increased response times in the compatible compared to the incompatible condition. The results are presented in Table 3.1 and Figure 3.1.

Table 3.1: Mean response times and standard deviations (in parentheses) as a function of Condition (Compatible vs Incompatible) and Group (Toward vs Away)

	Compatible	Incompatible
Toward	1300(363)msec	1144(409)msec
Away	1172(425)msec	1237(363)msec

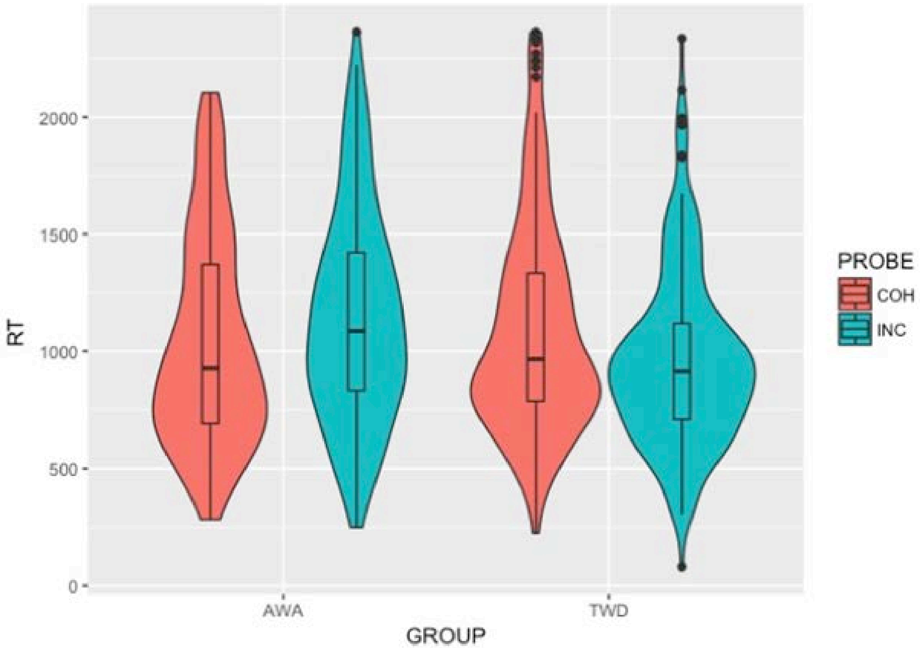


Figure 3.1: *Violin plots with medians and box plots as a function of Condition (Compatible vs Incompatible) and Group (Away vs Toward)*

3.3.2 ERP analysis

Visual examination of grand mean waveforms revealed no differences between conditions in any time windows when locked to verb onset. When time-locked to verb offset, there were no visible differences in the time windows associated with MP and RAP (-50-90 msec and 200-300 msec post-verb offset) (Figure 3.2). However, in the time window associated with the N400 (300-600 msec post verb offset), greater negativity was observed for Compatible compared to Incompatible trials. Statistical analyses confirmed these effects.

Linear mixed-effect regressions (lmer) were performed using the LmerTest package (Kuznetsova et al., 2017) implemented in R (R Core Team, 2017). The model included the sum-coded factors Group (Away vs Toward) and Condition (Compatible vs Incompatible) and their interaction. Participant and Item both included random intercepts and a random slope for Condition. Data points exceeding the mean plus or minus two standard deviations were considered outliers and removed (1% of the data). Below we report the results from the maximal random-effects structure (Barr et al., 2013). The model also met the criterion of having the smallest Akaike Information Criterion (AIC) (Matuschek, Kliegl, Vasishth, Baayen & Bates, 2017).

Motor response: MP and RAP

No significant differences emerged for the 4 midline electrodes entered into the model as a function of Condition, in either the MP time window (β=25,89, se = 27.27, t = 0.949, ns) or in the RAP time window (β=-40.59, se = 57.20, t = -1.015, ns).

Semantic treatment: 300-600 msec time window

The model involving 35 electrodes conducted on the data time-locked to verb offset revealed an effect of Condition (β= 2.59, se = 1.14, t= 2.27, p<.03), but not of Group (β= 2.28, se = 1.80, t= 1.27, ns) nor their interaction (β= -1.43, se = 1.37, t= -1.04, ns). Sentences describing actions that were compatible with the performed action elicited greater negativity in the waveform than incompatible sentences (Figures 3.2 and 3.3). A subsequent model, performed on the 9 central electrodes used in the time-frequency analysis (Fz, Cz, Pz, FC3, C3, CP3, FC4, C4, CP4) and in previous studies that examined motor-semantic interactions (Aravena et al., 2010; van Elk et al., 2010) revealed the same effect of Condition (β= -0.98, se = 0.47, t= -2.10, p<.04), and no interaction with Group (β= -0.35, se = -1.96, t= 0.35, ns).

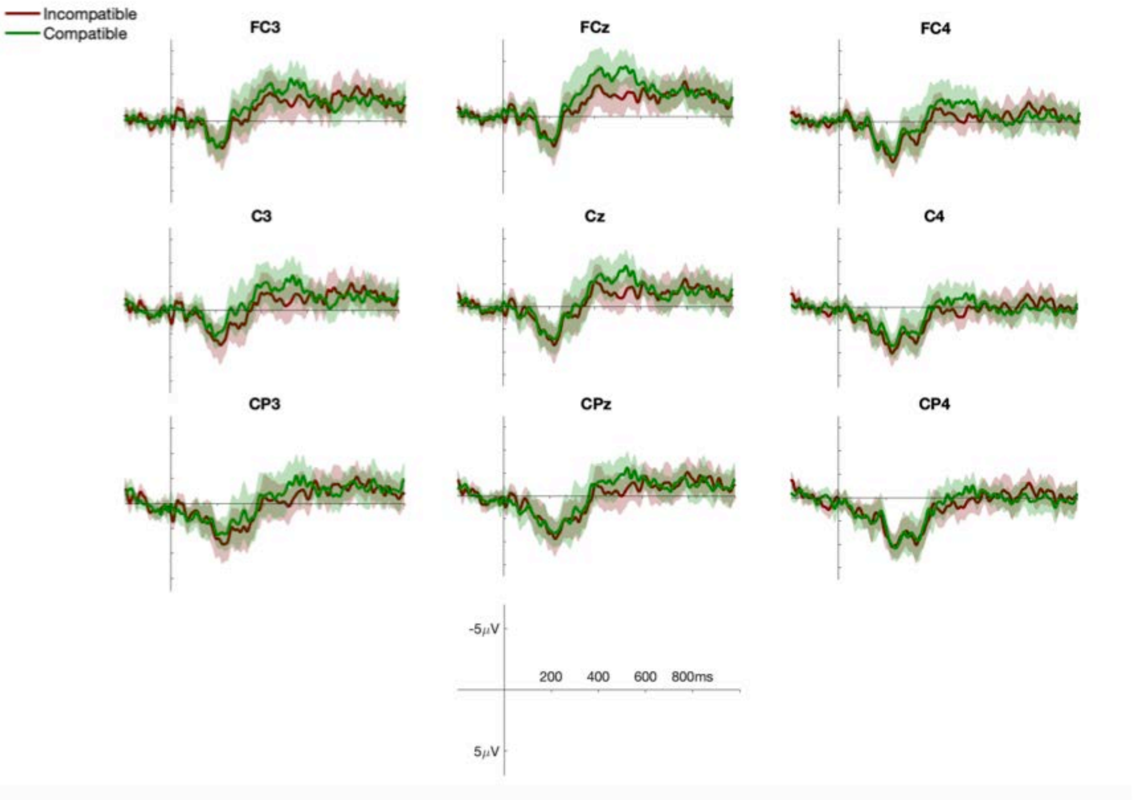


Figure 3.2: *Waveforms (and 95 % confidence intervals) for Incompatible (red) and Compatible (green) conditions with T0 corresponding to the verb offset*

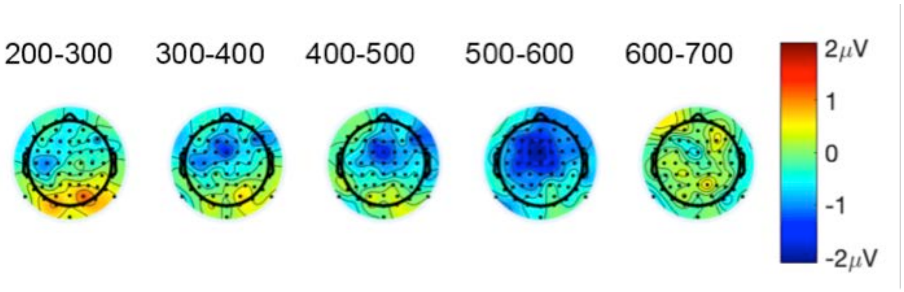


Figure 3.3: *Topographic maps of Compatible – Incompatible conditions plotted as a function of time over the post-stimulus interval following verb offset*

3.3.3 ERSP analysis

Compatible-Incompatible

Figure 3.4 shows ERSP for compatible and incompatible individually for 9 fronto-central and central electrodes, with T0 at verb onset. Mu-band activity was observed for central electrodes, most notably the C3 and Cz electrodes. Greater power in the mu-band was observed in the left-hemisphere, for both conditions. The mean verb offset, 490 msec, is indicated on the ERSP images. We observed that for the left hemisphere, the desynchronization in the mu-band emerged prior to verb offset whereas it emerged at or after verb offset in the right hemisphere.

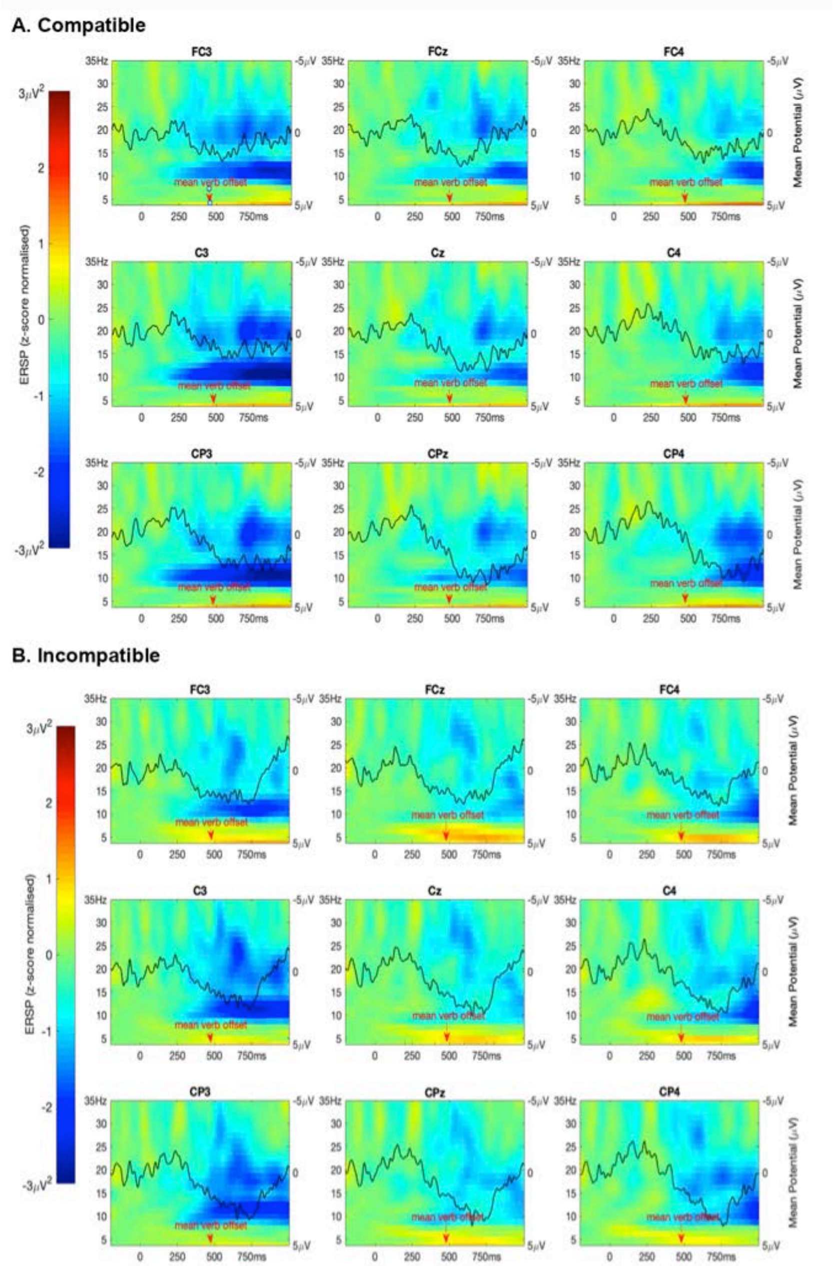
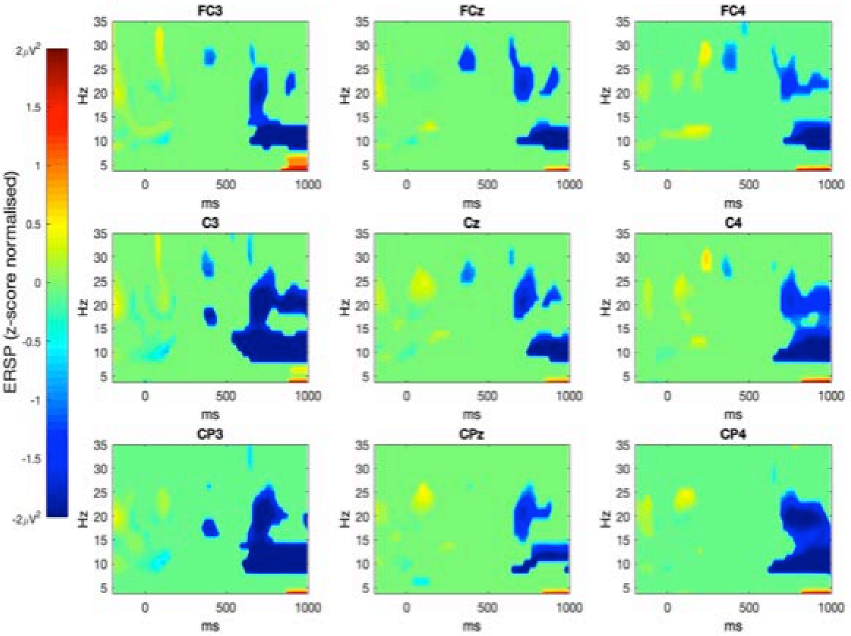


Figure 3.4: *Maps of ERSP for the Compatible (panel A) and Incompatible (panel B) conditions for 9 electrodes*

The post-stimulus ERSP was compared statistically to the baseline interval, preceding the verb onset, using a non-parametric permutation testing approach. The spectral power was z-score normalized in relation to the baseline interval (-200-0 msec). We corrected for multiple comparisons using FDR correction. This analysis was carried out for each electrode separately and included all frequencies of interest (4-35Hz). Significant ($p \leq .025$) ERSP was plotted and all non-significant activity was masked. Figure 3.5 presents the results of this test for the interval following the verb-onset. For both conditions we see significant ERD in the mu and beta bands that emerged in the latter part of the post-stimulus interval; the emergence of this ERD coincided roughly with the mean verb offset. This desynchronization in both mu and beta bands was present in both hemispheres. No differences in the mu-band emerged for comparisons between groups (Away vs Towards) or conditions (Compatible vs Incompatible).

A. Compatible



B. Incompatible

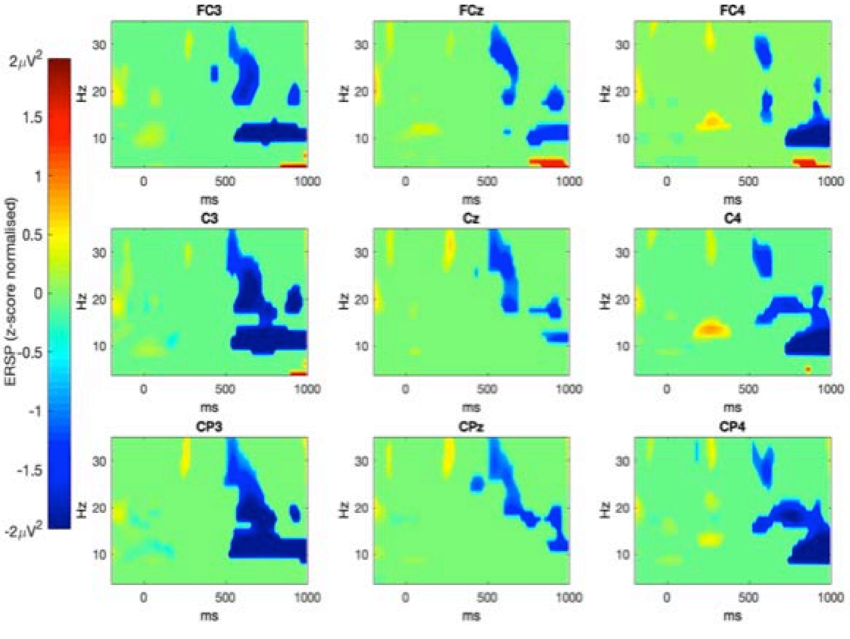


Figure 3.5: *Maps showing the significant ($p \leq .025$) time-frequency points for Compatible (A) and Incompatible (B) conditions*

Compatible and Incompatible Merged

Figure 3.6 presents the mean ERSP time-locked to verb onset over the 9 electrodes of interest (highlighted in the figure) for Compatible and Incompatible merged. The time point corresponding to the peak desynchronization for the mu-band (8-13Hz) and beta1 band (14-21Hz) was identified and a time window extending from 100 msec before to 100 msec after this peak time was defined for each frequency band. The topographies above the ERSP map present (left) the mean beta1 ERSP over a time window extending from 597 msec to 797 msec (peak desynchronization was identified at 697 msec) and (right) the mean mu-band ERSP over a time window from 850-1000 msec (peak occurring at 950 msec). A clear left-hemisphere bias can be seen for both mu and beta1 bands.

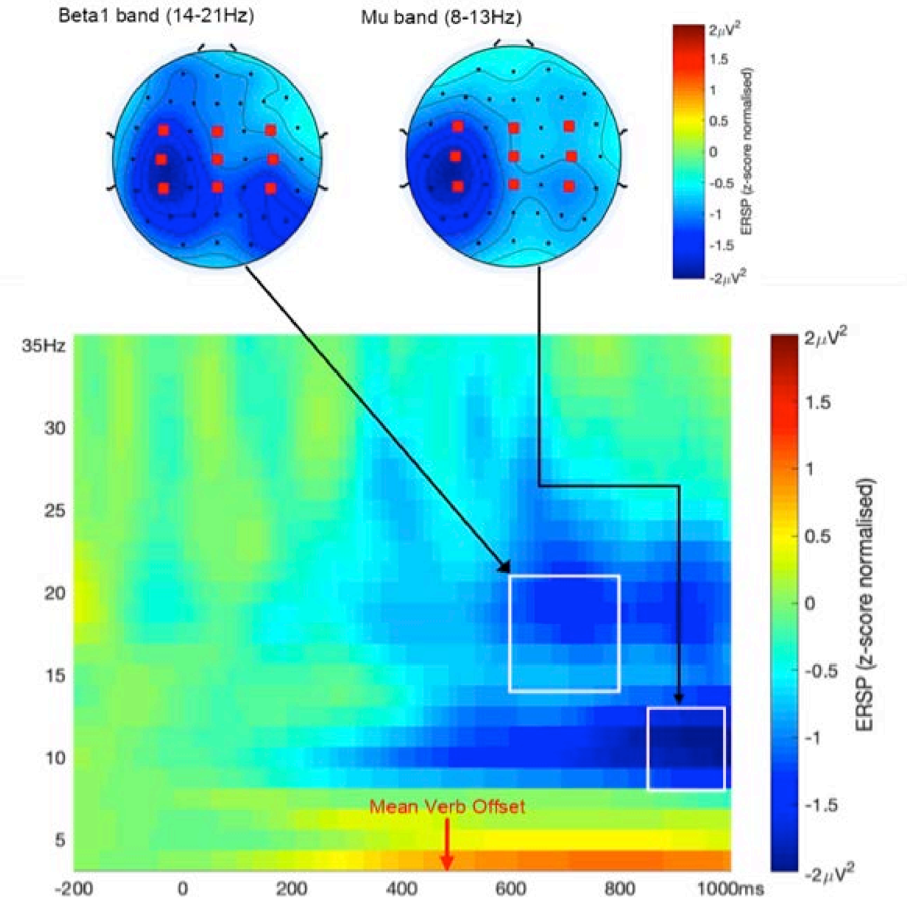


Figure 3.6: *The average ERSP over the 9 electrodes of interest (FC3, FCz, FC4, C3, Cz, C4, CP3, CPz, CP4) time-locked to verb onset*

3.4 Discussion

The current study aimed to explore the neural correlates of the interaction between motor action and language comprehension in order to determine whether and, if so, when motor preparation and semantic activation interact during linguistic processing. Specifically, we examined whether compatibility between action and action language produced bidirectional motor-semantic effects by measuring both behavioral and neurological responses. Based on previous results showing both facilitation and inhibition for action language compatibility, we performed ERP and time-frequency analyses on the EEG data to examine, first, the direction of effects and, second, whether they occurred in time windows associated with motor preparation and semantic processing. Time-frequency analyses revealed motor resonance starting around verb offset, in both conditions. No significant differences in either mu or beta ERD were found as a function of action language compatibility. The ERP results revealed clear interference of motor processes on semantic processing due to action language compatibility such that the N400 showed a greater negative deflection for compatible versus incompatible trials. In contrast, no earlier differences were found as a function of compatibility, i.e. in the time windows corresponding to the MP and RAP components. Hence, the combined pattern of time-frequency and ERP results showed that motor activation affected linguistic processing but that semantic processes did not affect motor preparation. Last, our behavioral response times do not provide any strong evidence of an ACE, in either direction.

Although most single-word studies time-lock their analysis to verb onset, in our study semantic processing appears to have been delayed such that a clear N1, P2, N400 pattern was visible at the offset of the critical verb. This can be attributed to the fact that participants had to understand complex sentences containing two main clauses, with the critical verb at the end of the second one, in past participle form. Accordingly, Moreno and colleagues found that motor activation occurred online during semantic integration across the sentence and not simply at the lexical level during single verb retrieval (Moreno et al., 2015). They interpreted their results as showing that the time course of mu desynchronization in sentence understanding underscores the time it takes for listeners or readers to converge the verb and noun(s) in order to form meaning.

3.4.1 ERPs

Semantic effects

Contrary to the results from previous studies using EEG to measure the neurophysiological correlates of the ACE, which showed an increased N400 for incompatible trials

(Aravena et al., 2010), we found a greater negative deflection for compatible compared to incompatible trials.

The direction of this N400 effect indicates that action language compatibility did not facilitate semantic processing, but interfered with it. These results are in line with behavioral studies showing that when semantic and motor processes overlap, this can cause inhibition (Buccino et al., 2005; Boulenger et al., 2006; de Vega et al, 2013). Indeed, in our study participants were asked to make a sensibility judgment at the end of an action sentence. Importantly, because the action verb was sentence-final, the sentence could only be fully understood, including the direction of the action described, at the very end of the sentence ("Suzanne a pris les cartes et les a distribuées" [Suzanne took the cards and distributed them]). Immediately following the sentence-final verb, participants saw an answer prompt. Our results suggest that semantic processing of the sentence occurred simultaneously with motor processing and created competition for processing resources. We posit that, during compatible trials, the motor preparation necessary for performing a compatible action competed with simulation of the action described in the sentence due to semantic processing, indicating shared neural substrates for action and semantic access. Such an effect supports the hypothesis that simulations can be involved in the understanding of action language (Barsalou, 1999).

Within the simulation framework, the Hand-Action-Network Dynamic Language Embodiment (HANDLE) model (García & Ibáñez, 2016) provides an interesting explanation for the two possible directions - facilitatory or inhibitory - of the effect of motor-language compatibility. The model proposes that accessing verbs denoting manual actions automatically activates neural networks that support motor action and processes. Processing hand-related expressions (HREs) produces, first, supra and later, sub-threshold activity levels. This determines whether motor resonance results in facilitation or interference for motor planning and semantic processing. According to the HANDLE model, if hand motor processes overlap or occur within 400 msec of the presentation of an HRE, competition for neural resources shared by two concurrent processes will cause interference. One of the processes will raise the activity above threshold level, interfering with the other process' access to the same resources, therefore either slowing down motor responses or hampering semantic processing. Although our behavioral results do not provide strong corroborating evidence of the inhibition of actual hand movements, our ERP results suggest that action planning did indeed interfere with semantic access.

Motor effects

In contrast to the robust effect found for semantic processing, we found no differences between conditions for a motor potentials (MP) and a re-afferent potentials (RAP). This result contrasts with that reported in Aravena et al.'s study (2010), in which compatible trials elicited larger MP amplitudes, an outcome interpreted as semantic priming of motor processing. It is important to note that in the Aravena and colleagues' study (2010), participants only had to indicate when they understood a sentence, using the same hand movement. Response mapping therefore likely occurred early in the sentence. Indeed, in traditional ACE studies, compatibility effects only occur when participants can plan movements early in sentence processing, as a result of a priming mechanism (Borreggine & Kaschak, 2006). According to Borreggine and Kaschak's (2006) feature binding account of the ACE, during language-induced simulation the feature (in our case the direction away or towards one's body) becomes bound to the simulation and is hence less available for movement execution, which cancels out motor priming. As mentioned above, in our experiment, given the sentence-final position of the verb, sentences could only be fully understood at the end; therefore, response mapping could only occur at that point, ruling out facilitation effects in motor preparation.

3.4.2 Behavioral effects

Much like previous studies that failed to replicate the ACE behaviorally (Papesh, 2015), our results do not show a clear behavioral ACE. Response times revealed an interaction such that participants who responded positively toward their body showed a trend to be hindered when the motor response matched the linguistic interpretation of the sentence whereas no significant difference emerged in the group of participants who responded away from the body. This interaction is not easily explained and does not match the pattern of results found for ERP components. Nonetheless, the trend for response times that was found specifically in the "Toward" group is coherent with the pattern of ERPs found for both groups, whereby responses were slower for compatible trials. Timing is known to have a strong influence on the direction of the behavioral ACE. As outlined above, the timing used in our study meant that participants accessed meaning at the end of the sentence, which could explain why we did not observe facilitation (Kaschak & Borreggine, 2008). On the other hand, the trend for an inhibition effect is in line with the predictions of the HANDLE model (García & Ibáñez, 2016), which assumes competition for common semantic-motor neural networks (García & Ibáñez, 2018).

The discrepancy between our behavioral results and those reported by Aravena et al. (2010) may be accounted for by the difference in task complexity. Although Aravena et al. used double-clause, verb-final sentences very similar to ours, in their experiment participants always answered with the same movement, i.e. they did not have to select a direction for answering. The movement was a simple button press and possibly required less executive control than was needed in our experiment, in which participants were requested to select an answer direction (“away” or “towards” their body) as a function of the semantic acceptability of the auditory sentences. As outlined by García and Ibáñez (2018), this could lead to a different relationship between motor preparation and semantic processing. Indeed, given that type of movement has been shown to influence the direction of the ACE (Shebani & Pulvermüller, 2018), it would stand to reason that indication of understanding versus judgment of semantic acceptability with differentiated answer directions could influence response time and hence the direction and occurrence of a behavioral ACE.

Discrepancies in timing of motor preparation during sentence treatment likely account for the inconsistency of the behavioral results across studies. Importantly, ours is not the first study to fail to find a solid behavioral ACE (cf. Papesh, 2015) and the literature clearly shows that robustness of the effect is contingent on a myriad of factors including timing, perspective, linguistic context, movement type and task (Alemanno et al., 2012; Aravena et al., 2012, 2014; Boulenger et al., 2006; Papeo et al., 2009; Shebani & Pulvermüller, 2018; Sato et al., 2008).

3.4.3 Time-frequency effects

In line with studies focusing on the time course of neural activity during language processing, we measured motor resonance as revealed by mu and beta desynchronization during action verb processing (Klepp et al., 2019; Moreno, Vega & León, 2013; Schaller, Weiss & Müller, 2017; van Elk et al., 2010). Studies focusing on motor activation as a result of single verb processing generally link mu and beta desynchronization to facilitated action language processing (Klepp et al., 2019; Niccolai et al., 2014; Zappa et al. 2019). In sentence processing studies, although the pattern of mu and beta desynchronization has sometimes been interpreted as showing increased effort in semantic processing (Lam et al., 2017), most studies interpret desynchronization in these two bands as showing motor resonance (Moreno et al., 2015; Schaller et al., 2017; van Elk et al., 2010).

In our study, we went beyond passive listening and reading tasks used in the above studies and manipulated action compatibility while measuring motor resonance during action sentence processing. We observed ERD in the mu and beta bands, for both incompatible and compatible trials, starting around the offset of the verbs, hence late in the auditory stream, in line with recent studies (Moreno et al., 2015) and throughout the window during which ERPs showed semantic processing, i.e. following verb offset. No differences emerged as a function of condition. Mu and beta desynchronization have been linked to both motor resonance that results from simulation during action language processing (Moreno et al., 2015; Niccolai et al., 2014; van Elk et al., 2010) and motor preparation and observation (Caetano et al., 2007; Koelewijn et al., 2008; Pfurtscheller & Lopes da Silva, 1999; Pineda, 2005). Given that participants were both processing action verbs and preparing a movement to answer in the time window during which we observe mu and beta desynchronization, we cannot claim that these effects are due to one or the other process alone. We hypothesize that while the mu and beta desynchronization found during verb processing are likely partially linked to motor resonance from simulation, it may also result from motor preparation for the upcoming response. This is in line with results reported by Schaller et al. (2017) who also found beta ERD in relation to the processing of action verbs but roughly 1200 msec after verb onset and close to the actual motor response. Schaller et al. (2017) also failed to find a difference in the degree of beta ERD as a function of the interpretation of the verb, i.e. whether a concrete or abstract action (i.e. to grasp a hand vs an idea).

3.4.4 Conclusion

Embodied cognition grounds cognitive processes in sensory-motor experiences, and embodied semantics claims that action language processing involves the simulation of such experiences. The present study provides cortical evidence of concurrent activation of sensory-motor networks during the linguistic processing of action verbs. Difficulty in semantic processing, as revealed by the N400, emerged as a result of compatibility between action language and motor planning, suggesting that motor and semantic processes competed for neural resources. Our study adds to the literature that teases apart facilitatory and inhibitory ACE effects in an attempt to better understand the common mechanisms underlying motor and linguistic processes. Importantly, whereas the majority of ACE studies focus on semantic-to-motor effects, we provide an interesting set of results showing motor-to-semantic effects.

3.5 Annex

Away Sentences

Suzanne a pris les cartes et les a distribuées.
Michel a pris des brochures et les a distribuées.
Céline a sorti sa carte bleue et a payé.
Bastien a sorti un billet et a payé.
Lise a vu le vase sur le frigo et l'a attrapé.
Georges a vu la boîte dans le placard et l'a attrapée.
Jacques a essayé d'attraper sa soeur mais elle l'a repoussé.
Lola voulait jouer avec son frère mais il l'a repoussée.
Visant le bras de l'homme, Anna l'a frappé.
Enervé contre son adversaire, le boxeur l'a frappé.
Maude a ouvert le champagne et l'a servi.
Serge a pris de l'eau pour les sportifs et les a servis.
Comme le tiroir était ouvert, Noël l'a fermé.
Comme le placard était ouvert, Laure l'a fermé.
Elsa a pris le ballon et l'a jeté.
Antoine a pris un caillou et l'a jeté.
Le lait étant sorti du frigo, Paul l'y a remis.
La tarte n'étant pas cuite, Claude l'a remise à cuire.
Diane a soulevé le tableau et l'a accroché.
Rémi a pris le tableau et l'a accroché.

Toward Sentences

Arthur a trouvé le tiroir, puis l'a ouvert.
Devant le placard, Thérèse l'a ouvert.
Stéphanie a pris un raisin et l'a mangé.
Patrick a coupé une tranche de pain et l'a mangée.
Ayant la bouche sale, Elise s'est essuyée.
Matthieu a pris sa serviette et s'est essuyé.
Camille a pris sa couette et s'est couverte.
Olivier a pris une couverture et s'est couvert.
Une fois le billet sorti, Gérard il l'a récupéré.
Quand l'argent est sorti, Sophie l'a récupéré.

Margot a sorti un mouchoir et s’est mouchée.
Cyril a pris un mouchoir et s’est mouché.
L’acteur a récupéré son chapeau et l’a mis.
Marie a pris ses lunettes, et les a mises.
Gabriel s’est servi une tasse de café et l’a bu.
Emilie a pris son verre de vin l’a bu.
Quand le téléphone a sonné, Rodolf a répondu.
Le téléphone a sonné et Christine a répondu.
Théo a pris un rasoir et s’est rasé.
Leo a pris son rasoir et s’est rasé.

Anomalous Sentences

Jules a pris ses neurones et les a distribués.
Comme Luis n’avait plus d’argent, il a tout payé.
Isabelle a vu feu et l’a attrapé.
Edouard aimait tellement sa copine qu’il l’a repoussée.
Pensant faire plaisir à son ami, Julie l’a frappé.
Comme le patient ne pouvait rien manger, l’infirmière l’a servi.
Voulant montrer son cahier à ses parents, Stéphane l’a fermé.
Pour ranger, Johan a jeté des verres par terre.
Quand l’eau s’est évaporée, Oscar l’a remise dans le verre.
Daniel a pris la soupe et l’a accrochée.
Voulant se cacher, Christian a ouvert la porte.
Monique a pris un tas de papier et l’a mangé.
Gilles a ouvert son ordinateur et s’est essuyé avec.
Comme Pascal avait très chaud, il s’est vite couvert.
Une fois que sa fille était partie, Jean-Luc l’a récupérée.
En plein sommeil, Nadine s’est mouchée.
Ayant chaud, Aurélie a mis son manteau.
André s’est coupé une part de pizza et l’a bue.
Ne voulant pas parler, Véronique a répondu au téléphone.
Le bambin s’est rasé devant la glace.

References

Alemanno, F., Houdayer, E., Cursi, M., Velikova, S., Tettamanti, M., Comi, G., Cappa, S. F., & Leocani, L. (2012). Action-related semantic content and negation polarity modulate motor areas during sentence reading: An event-related desynchronization study. *Brain Research*, *1484*, 39–49.

Aravena, P., Courson, M., Frak, V., Cheylus, A., Paulignan, Y., Deprez, V., & Nazir, T. (2014). Action relevance in linguistic context drives word-induced motor activity. *Frontiers in Human Neuroscience*, *8*.

Aravena, P., Delevoye-Turrell, Y., Deprez, V., Cheylus, A., Paulignan, Y., Frak, V., & Nazir, T. (2012). Grip Force Reveals the Context Sensitivity of Language-Induced Motor Activity during "Action Words" Processing: Evidence from Sentential Negation. *PLOS ONE*, *7*(12), e50287.

Aravena, P., Hurtado, E., Riveros, R., Cardona, J. F., Manes, F., & Ibáñez, A. (2010). Applauding with closed hands: Neural signature of action-sentence compatibility effects. *PloS One*, *5*(7), e11751.

Aziz-Zadeh, L., Wilson, S. M., Rizzolatti, G., & Iacoboni, M. (2006). Congruent Embodied Representations for Visually Presented Actions and Linguistic Phrases Describing Actions. *Current Biology*, *16*(18), 1818–1823.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Barsalou, L. W. (1999). Perceptual symbol systems. Behavioral and Brain Sciences , 22 (4), 577-660.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Bigi, B. (2015). SPPAS - Multi-lingual approaches to the automatic annotation of speech. *The Phonetician*, *111–112*(ISSN:0741-6164), 54–69.

Black, J. B., Khan, S. A., & Huang, S.-C. D. (2014). Video and Computer Games as Grounding Experiences for Learning. In F. C. Blumberg (Ed.), *Learning by Playing* (pp. 290–301). Oxford University Press.

Borreggine, K. L., & Kaschak, M. P. (2006). The Action-Sentence Compatibility Effect: It's All in the Timing. *Cognitive Science*, *30*(6), 1097–1112.

Boulenger, V., Hauk, O., & Pulvermüller, F. (2009). Grasping ideas with the motor system: Semantic somatotopy in idiom comprehension. *Cerebral Cortex (New York, N.Y.: 1991)*, *19*(8), 1905–1914.

Boulenger, V., Roy, A., Paulignan, Y., Deprez, V., Jeannerod, M., & Nazir, T. (2006). Cross-talk between language processes and overt motor behavior in the first 200 msec of processing. *Journal of Cognitive Neuroscience*.

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

Caetano, G., Jousmäki, V., & Hari, R. (2007). Actor's and observer's primary motor cortices stabilize similarly after seen or heard motor actions. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, *104*(21), 9058–9062.

Chwilla, D. J., Kolk, H. H. J., & Vissers, C. T. W. M. (2007). Immediate integration of novel meanings: N400 support for an embodied view of language comprehension. *Brain Research*, *1183*, 109–123.

Collins, J., Pecher, D., Zeelenberg, R., & Coulson, S. (2011). Modality Switching in a Property Verification Task: An ERP Study of What Happens When Candles Flicker after High Heels Click. *Frontiers in Psychology*, *2*, 10.

de Vega, M., Moreno, V., & Castillo, D. (2013). The comprehension of action-related sentences may cause interference rather than facilitation on matching actions. *Psychological Research*, *77*(1), 20–30.

de Vega, M., & Urrutia, M. (2011). Counterfactual sentences activate embodied meaning: An action–sentence compatibility effect study. *Journal of Cognitive Psychology*, *23*(8), 962–973.

de Vega, M., Morera, Y., León, I., Beltrán, D., Casado, P., & Martín-Loeches, M. (2016). Sentential Negation Might Share Neurophysiological Mechanisms with Action Inhibition. Evidence from Frontal Theta Rhythm. *Journal of Neuroscience*, *36*(22), 6002–6010.

Diefenbach, C., Rieger, M., Massen, C., & Prinz, W. (2013). Action-Sentence Compatibility: The Role of Action Effects and Timing. *Frontiers in Psychology*, *4*.

Dudschig, C., de la Vega, I., & Kaup, B. (2014). Embodiment and second-language: Automatic activation of motor responses during processing spatially associated L2 words and emotion L2 words in a vertical Stroop paradigm. *Brain and Language*, *132*, 14–21.

Fargier, R., Paulignan, Y., Boulenger, V., Monaghan, P., Reboul, A., & Nazir, T. A. (2012). Learning to associate novel words with motor actions: Language-induced motor activity following short training. *Cortex*, *48*(7), 888–899.

García, A. M., & Ibáñez, A. (2016). A touch with words: Dynamic synergies between manual actions and language. *Neuroscience and Biobehavioral Reviews*, *68*, 59–95.

García, A. M., & Ibáñez, A. (2018). When embodiment breaks down: Language deficits as novel avenues into movement disorders. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *100*, 1–7.

Gianelli, C., Farnè, A., Salemme, R., Jeannerod, M., & Roy, A. C. (2011). The Agent is Right: When Motor Embodied Cognition is Space-Dependent. *PLoS ONE*, *6*(9).

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, *9*(3), 558–565.

Glenberg, A. M., Sato, M., Cattaneo, L., Riggio, L., Palumbo, D., & Buccino, G. (2008). Processing abstract language modulates motor system activity. *Quarterly Journal of Experimental Psychology (2006)*, *61*(6), 905–919.

Grandchamp, R., & Delorme, A. (2011). Single-trial normalization for event-related spectral decomposition reduces sensitivity to noisy trials. *Frontiers in Psychology*, *2*, 236.

Hald, L. A., Marshall, J.-A., Janssen, D. P., & Garnham, A. (2011). Switching Modalities in A Sentence Verification Task: ERP Evidence for Embodied Language Processing. *Frontiers in Psychology*, *2*, 45.

Hauk, O., & Tschentscher, N. (2013). The body of evidence: What can neuroscience tell us about embodied semantics? *Frontiers in Psychology*, *4*.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic Representation of Action Words in Human Motor and Premotor Cortex. *Neuron*, *41*(2), 301–307.

Hobson, H., & Bishop, D. V. M. (2016). Mu suppression – A good measure of the human mirror neuron system? *Cortex*.

Horoufchin, H., Bzdok, D., Buccino, G., Borghi, A. M., & Binkofski, F. (2018). Action and object words are differentially anchored in the sensory motor system—A perspective on cognitive embodiment. *Scientific Reports*, *8*(1), 6583.

Kaschak, M. P., & Borreggine, K. L. (2008). Temporal dynamics of the action-sentence compatibility effect. *Quarterly Journal of Experimental Psychology (2006)*, *61*(6), 883–895.

Klepp, A., Dijk, H. van, Niccolai, V., Schnitzler, A., & Biermann-Ruben, K. (2019). Action verb processing specifically modulates motor behaviour and sensorimotor neuronal oscillations. *Scientific Reports*, *9*(1), 1–11.

Koelewijn, T., van Schie, H. T., Bekkering, H., Oostenveld, R., & Jensen, O. (2008). Motor-cortical beta oscillations are modulated by correctness of observed action. *NeuroImage*, *40*(2), 767–775.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*(12), 463–470.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology*, *62*, 621–647.

Kutas, M., & Hillyard, S. A. (1980). Event-related brain potentials to semantically inappropriate and surprisingly large words. *Biological Psychology*, *11*(2), 99–116.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13).

Lam, K. J. Y., Bastiaansen, M. C. M., Dijkstra, T., & Rueschemeyer, S.-A. (2017). Making sense: Motor activation and action plausibility during sentence processing. *Language, Cognition and Neuroscience*, *32*(5), 590–600.

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris*, *102*(1), 59–70.

Makeig, S. (1993). Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones. *Electroencephalography & Clinical Neurophysiology*, *86*(4), 283–293.

Makeig, S., Debener, S., Onton, J., & Delorme, A. (2004). Mining event-related brain dynamics. *Trends in Cognitive Sciences*, *8*(5), 204–210.

Matsumoto, M., Okada, E., Ichihara, D., Chiba, K., Toyama, Y., Fujiwara, H., Momoshima, S., Nishiwaki, Y., Hashimoto, T., Inoue, T., Watanabe, M., & Takahata, T. (2010). Prospective ten-year follow-up study comparing patients with whiplash-associated disorders and asymptomatic subjects using magnetic resonance imaging. *Spine*, *35*(18), 1684–1690.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315.

Moreno, I., de Vega, M., León, I., Bastiaansen, M. C. M., Lewis, A. G., & Magyari, L. (2015). Brain dynamics in the comprehension of action-related language: A time frequency analysis of mu rhythms. *NeuroImage*, *109*(April), 50–62.

Moreno, I., de Vega, M., & León, I. (2013). Understanding action language modulates oscillatory mu and beta rhythms in the same way as observing actions. *Brain and Cognition*, *82*(3), 236–242.

Niccolai, V., Klepp, A., Weissler, H., Hoogenboom, N., Schnitzler, A., & Biermann-Ruben, K. (2014). Grasping hand verbs: Oscillatory beta and alpha correlates of action-word processing. *PloS One*, *9*(9), e108059.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2010). *FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data* [Research Article]. Computational Intelligence and Neuroscience; Hindawi.

Papeo, L., Vallesi, A., Isaja, A., & Rumiati, R. I. (2009). Effects of TMS on different stages of motor and non-motor verb processing in the primary motor cortex. *PloS One*, *4*(2), e4508.

Papesh, M. H. (2015). Just out of reach: On the reliability of the action-sentence compatibility effect. *Journal of Experimental Psychology. General*, *144*(6), e116-141.

Pfurtscheller, G., & Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: Basic principles. *Clinical Neurophysiology*, *110*(11), 1842–1857.

Pineda, J. A. (2005). The functional significance of mu rhythms: Translating “seeing” and “hearing” into “doing.” *Brain Research. Brain Research Reviews*, *50*(1), 57–68.

Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences*, *22*(2), 253–336.

Pulvermüller, F. (2002). A brain perspective on language mechanisms: From discrete neuronal ensembles to serial order. *Progress in Neurobiology*, *67*(2), 85–111.

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews. Neuroscience*, *6*(7), 576–582.

Pulvermüller, F. (2012). Meaning and the brain: The neurosemantics of referential, interactive, and combinatorial knowledge. *Journal of Neurolinguistics*, *25*(5), 423–459.

Pulvermüller, F., & Fadiga, L. (2010). Active perception: Sensorimotor circuits as a cortical basis for language. *Nature Reviews. Neuroscience*, *11*(5), 351–360.

Raposo, A., Moss, H. E., Stamatakis, E. A., & Tyler, L. K. (2009). Modulation of motor and premotor cortices by actions, action words and action sentences. *Neuropsychologia*, *47*(2), 388–396.

Sato, M., Mengarelli, M., Riggio, L., Gallese, V., & Buccino, G. (2008). Task related modulation of the motor system during language processing. *Brain and Language*, *105*(2), 83–90.

Schaller, F., Weiss, S., & Müller, H. M. (2017). "Pushing the Button While Pushing the Argument": Motor Priming of Abstract Action Language. *Cognitive Science*, *41*(5), 1328–1349.

Shebani, Z., & Pulvermüller, F. (2018). Flexibility in Language Action Interaction: The Influence of Movement Type. *Frontiers in Human Neuroscience*, *12*.

Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., Fazio, F., Rizzolatti, G., Cappa, S. F., & Perani, D. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience*, *17*(2), 273–281.

Thompson-Schill, S. L. (2003). Neuroimaging studies of semantic memory: Inferring "how" from "where." *Neuropsychologia*, *41*(3), 280–292.

van Elk, M., van Schie, H. T., Zwaan, R. A., & Bekkering, H. (2010). The functional role of motor activation in language processing: Motor cortical oscillations support lexical-semantic retrieval. *NeuroImage*, *50*(2), 665–677.

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

Zappa, A., Bolger, D., Pergandi, J.-M., Mallet, P., Dubarry, A.-S., Mestre, D., & Frenck-Mestre, C. (2019). Motor resonance during linguistic processing as shown by EEG in a naturalistic VR environment. *Brain and Cognition*, *134*, 44–57.

Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: Motor resonance in language comprehension. *Journal of Experimental Psychology: General*, *135*(1), 1–11.

Chapter 4

Motor Resonance during linguistic processing as shown by EEG in a naturalistic VR environment

Zappa, A., Bolger, D., Pergandi, J.-M., Mallet, P., Dubarry, A.-S, Mestre, D. & Frenck-Mestre, C., *Brain and Cognition*, 2019

Brain and Cognition 134 (2019) 44–57

Contents lists available at ScienceDirect

Brain and Cognition

journal homepage: www.elsevier.com/locate/b&c

Motor resonance during linguistic processing as shown by EEG in a naturalistic VR environment

Ana Zappa[a,e,d,f,*], Deirdre Bolger[a,e,f,d], Jean-Marie Pergandi[a,b,c], Pierre Mallet[a,b,c], Anne-Sophie Dubarry[a,d,e], Daniel Mestre[a,b,c,d,f], Cheryl Frenck-Mestre[a,d,e,f]

[a] *Aix-Marseille Univ, France*
[b] *Institute of Movement Sciences, France*
[c] *Mediterranean Virtual Reality Center, France*
[d] *Centre National de Recherche Scientifique, France*
[e] *Laboratoire Parole et Langage, France*
[f] *Institute of Language, Communication and the Brain, France*

ABSTRACT

Embodied cognition studies have shown motor resonance during action language processing, indicating that linguistic representations are at least partially multimodal. However, constraints of this activation linked to linguistic and extra-linguistic context, function and timing have not yet been fully explored. Importantly, embodied cognition binds social and physical contexts to cognition, suggesting that more ecologically valid contexts will yield more valid measures of cognitive processing. Herein, we measured cortical motor activation during language processing in a fully immersive Cave automatic virtual environment (CAVE). EEG was recorded while participants engaged in a Go/No-Go task. They heard action verbs and, for Go trials, performed a corresponding action on a virtual object. ERSP (event-related spectral perturbation) was calculated during verb processing, corresponding to the pattern of power suppression (event-related desynchronization – ERD) and enhancement (event-related synchronization – ERS) relative to the reference interval. Significant ERD emerged during verb processing in both the μ (8–13 Hz) and beta band (20–30 Hz) for both Go and No-Go trials. μ *ERD* emerged in the 400–500 msec time window, associated with lexical-semantic processing. Greater μ ERD emerged for Go compared to No-Go trials. The present results provide compelling evidence in a naturalistic setting of how motor and linguistic processes interact.

1. Introduction

A currently debated topic in cognitive psychology is the involvement of motor processes in language processing. Studies that approach language from an embodied cognition perspective have produced evidence that language comprehension involves perceptual and motor systems, indicating that linguistic representations are either partially or completely multimodal (Barsalou, 1999, 2008; Wilson & Golonka, 2013). Both neuroimaging and behavioral results have pointed to the recruitment of sensorimotor systems during semantic access and the overlap of these processes has often been interpreted as evidence that one performs mental simulations of situations to understand language (Barsalou, 1999; Glenberg & Kaschak, 2002; Pulvermüller, 2005). However, classical models of language comprehension posit that language representation is amodal and independent of perceptual and motor systems, suggesting that these motor activations are post-lexical and do not play a causal role in language processing (Fodor, 1980, 1987; Mahon & Caramazza, 2008). Moreover, recent studies have shown that motor activation during language processing can be modified or neutralized by changes in linguistic context, task and timing (Aravena et al., 2012; Boulenger et al., 2006; Sato, Mengarelli, Riggio, Gallese, & Buccino, 2008). Hence, the implications of the involvement of motor processes in language comprehension are still not fully understood. Importantly, studies focusing on motor and linguistic interactions have yet to use set-ups that take into account the multimodality of language, to provide a closer-to-life experience under which to observe the bilateral influence between action and cognition (Peeters, 2019). The present study sought to fill this gap by placing participants in a three-dimensional virtual environment in which they manipulated virtual objects in response to linguistic cues. This allowed us to examine the hypothesis that motor representations are part and parcel of the linguistic representation of action verbs (Aravena et al., 2012; Pulvermüller, 2005) in a realistic environment, as opposed to the impoverished conditions that are most often used. The recording of EEG allowed us to clearly examine the neurological signature of motor planning (Aravena et al., 2010; Funderud et al., 2012) as well as any evidence of motor activation during linguistic processing.

---

[*] Corresponding author.
E-mail address: ana.zappa@lpl-aix.fr (A. Zappa).

https://doi.org/10.1016/j.bandc.2019.05.003
Received 18 February 2019; Received in revised form 1 May 2019; Accepted 6 May 2019

1.1. Behavioral evidence of motor involvement in linguistic processing

One way to observe the influence of motor processes on language comprehension is to manipulate motor activation through movement or motor planning and observe its influence on lexico-semantic processing. Behavioral studies using an Action-Sentence compatibility effect paradigm (ACE) have shown that compatibility between action language and the movement needed to produce a manual response can either hinder or facilitate response times depending upon timing. In semantic decision tasks, a negative ACE is found, i.e. the inhibition of response times for a compatible movement, within the first 400 msec during or after single action verb processing (Sato et al., 2008; Spadacenta, Gallese, Fragola, & Mirabella, 2014). When sentence comprehension is required, a negative ACE occurs within 700 msec post onset of the critical verb (García & Ibáñez, 2016, but see Repetto, Cipresso, & Riva, 2015). However, if response planning occurs at sentence onset *or* if movement onset is delayed to after the single verb or action sentence is fully processed, a positive ACE is obtained (Boulenger et al., 2006; de Vega, Moreno, & Castillo, 2013; Diefenbach, Rieger, Massen, & Prinz, 2013; Kaschak & Borreggine, 2008; for a review see García & Ibáñez, 2016). These results suggest that interference occurs when motor and linguistic processes overlap temporally due to competition for shared neural resources. Facilitation, on the other hand, seems to be due to a priming effect, upstream of competition but nonetheless indicating shared neural resources. Given that most ACE studies reveal effects late in sentence comprehension (700–2200 msec post stimulus), they do not rule out the possibility that the ACE effect is caused by post-lexical motor imagery (Toni, de Lange, Noordzij, & Hagoort, 2008).

1.2. Electrophysiological evidence

To overcome the limitations of behavioral measures as concerns the onset of motor influences on linguistic processing, several studies have adopted the recording of electrophysiological responses. Aravena and colleagues focused on the precise timing of motor-semantic effects using EEG to examine early motor-related ERPs (Readiness Potential (RP)) as well as later, linguistically related ERPs (N400) in an ACE paradigm (Aravena et al., 2010). Participants listened to action sentences while performing congruent and incongruent actions. Congruent actions led to an increase in the magnitude of the RP, revealing an effect of linguistic processing on motor preparation. In addition, in comparison to congruent trials, incongruent actions elicited a greater N400-like response, thus revealing the interference of motor execution in semantic processing. This pattern of results was interpreted as illustrating a robust ACE which, crucially, could not be attributed to post-lexical effects but supports the hypothesis of early interactions between sensorimotor and semantic processing. In the current study, we built upon these results first, by using a virtual environment in which participants performed movements actually related to the action verbs as opposed to a rather impoverished environment or only imagined movement (cf. Peeters, 2019). Second, we examined motor activation during language processing as revealed by time-frequency analyses as opposed to ERP components alone. As outlined below, while several studies have used time-frequency analyses to quantify motor cortex activity proper, this approach has been applied to the study of embodied cognition less frequently (cf. Fargier et al., 2012, for a discussion).

Numerous studies have used EEG to quantify motor cortex activity by measuring oscillatory activity via time-frequency decomposition of the EEG signal to examine the pattern of cortical response to motor planning. More recently, post-stimulus spectral estimation methods have been applied to study the pattern of event-related desynchronization (ERD) and synchronization (ERS), which corresponds to power suppression and enhancement respectively, in the time interval following stimulus onset compared to a pre-stimulus baseline. One method, event-related spectral perturbation (ERSP) (Grandchamp & Delorme, 2011; Makeig, 1993; Makeig, Debener, Onton, & Delorme, 2004), which we applied in the present study, groups ERD and ERS and calculates the power spectrum on a trial-by-trial basis. This approach can be applied to study the cortical response to motor events (for a review see Hobson & Bishop, 2016). Specifically, mu, or μ (8–13 Hz), and beta (13–30 Hz) rhythms are synchronized patterns of electrical activity recorded over the sensorimotor cortex whose suppression is associated with performing and observing movement (Caetano, Jousmäki, & Hari, 2007; Koelewijn, van Schie, Bekkering, Oostenveld, & Jensen, 2008; Pfurtscheller & Lopes da Silva, 1999; Pineda, 2005). A decrease in the μ rhythm has also been linked to motor imagery (Matsumoto et al., 2010).

In addition to movement proper, various recent studies have shown μ rhythm ERD as a function of action language processing. Language studies using time-frequency analysis have found μ and beta ERD for action-related sentences, directly showing motor resonance *during* the retrieval of lexical-semantic information as opposed to post-lexical retrieval of kinematic imagery (van Elk, van Schie, Zwaan, & Bekkering, 2010). Reading single verbs related to the body caused soma topical μ ERD (von Nicolai et al., 2014). Moreover, greater μ rhythm suppression has been observed for action language compared to abstract language (Alemanno et al., 2012; Moreno et al., 2015). To examine how the acquisition of lexical items might be directly influenced by the motor system, Fargier et al. (2012) conducted a learning paradigm using EEG. Participants learned novel words, either in association with motor actions or with abstract animated images. After two training sessions on a first day of learning, participants showed greater μ suppression while processing words learned in the motor action condition compared to the control condition. Nonetheless, based on the distribution of their effects across 2 days of training, Fargier et al. (2012) argued that the cortical regions conjointly activated by motor and linguistic processing were confined to convergence areas (i.e. more frontal regions as opposed to the central parietal areas assumed to subserve sensorimotor activity). Moreover, although these studies have all used time-locked μ suppression as a marker of motor neuron activity, it is important to keep in mind that the μ frequency band (8–13 Hz) overlaps with the alpha frequency band (8–12 Hz), which is reflective of attentional fluctuation, and it has been claimed that the two are often confounded (Hobson & Bishop, 2016). On the other hand, several researchers working with μ band oscillations argue that they reflect neural activity in the motor and premotor cortex and can be measured in fronto-central sites (but see Fargier et al., 2012) versus occipital sites for the alpha band (Moreno et al., 2015). Therefore, a consensus in terms of how to distinguish between these two has not been reached.

1.3. Virtual reality as a tool to study embodiment

A caveat of investigating motor activation during language processing, especially when working within an embodied cognition framework, is linked to the physical and environmental limitations imposed by neurolinguistic study protocols. Laboratory experimental tasks are generally performed in isolated and decontextualized environments, due to the need to control variables that could influence participants' responses (Peeters, 2019). These experiments often use single words or sentences presented in isolation, along with, if any, simplistic visual information on a computer screen. This very decontextualization could in turn affect how language is processed. Indeed, real-world language processing generally occurs in much richer environments and, importantly, interlocutors, social context and physical cues have a strong influence on how language is understood (Knoeferle, 2015). In the last two decades, the discrepancy between real-life language processing and that which takes place in an experimental environment has been brought to light in the context of embodied cognition (Tromp, Peeters, Meyer, & Hagoort, 2018), according to which our bodily states and actions are heavily implicated in how we communicate and process information (Atkinson, 2010). When we communicate, speech and gesture systems interact to convey and comprehend meaning; the two systems have been suggested to comprise an integrated system (Goldin-Meadow, 2011; Graziano & Gullberg, 2018). Therefore, the more real-world and situated the language processing environment is, the more physically implicated and natural participants will feel and more applicable the results obtained will be to real-life processing (Peeters, 2019). Modern technological advances render possible the study of cognitive processes in their actual contexts (Ladouce, Donaldson, Dudchenko, & Ietswaart, 2017). Notably, in the present study we capitalized on a CAVE Automatic Virtual Environment (CAVE) system to examine the interaction of language and motor processes during the lexical access of auditory verbs, in the aim of testing the hypothesis that motor representations are part and parcel of the linguistic representations of action verbs (Pulvermüller, 1999). Several studies of language acquisition in adults have demonstrated clear benefits of a virtual environment (Legault et al., 2019; Repetto, Colombo, & Riva, 2015). As concerns the interaction of linguistic and motor processing, Repetto, Cipresso et al. (2015) found that simulating actions in a virtual environment facilitated the semantic processing of action verbs that involved the same effector (i.e. simulated running in a virtual park facilitated processing of verbs entailing movement of the foot). Simply viewing the motion did not influence semantic processing. Hence, virtual motion can elicit stimulation of the motor system, which can in turn affect linguistic processing of overlapping information.

Virtual reality (VR) experimental paradigms have indeed gained popularity as they offer a more ecological and yet controlled environment in which to test a wide variety of phenomena including language processing (Peeters, 2019; Repetto, Cipresso et al., 2015; Tromp et al., 2018). VR paradigms consist of 3-D environments that provide participants with visual and auditory stimuli, while allowing them to interact and receive real-time feedback from their actions via a graphic rendering system. Participants' movements are often tracked and recorded using input tools (trackers, gloves, a mouse or joystick) (Burdea & Coiffet, 2003). Three basic types of VR environments exist: a computer monitor, a head-mounted display (HMD) and the CAVE Automatic Virtual Environment (CAVE) system. Computer monitors are considered non-immersive due to the small percentage of the participant’s visual field that they occupy (Repetto, 2014). HMDs provide an immersive experience. One interesting aspect of HMDs is that, contrary to a “classical” 3-sided CAVE, they visually isolate the participant from the real world. However, the downside is that participants no longer see their own bodies. This commonly leads to a sensation of self-flotation. Indeed, we carried out experiments comparing CAVE and HMD in a simple spatial task, i.e. walking through an aperture (Mestre, Louison, & Ferlay, 2016). We found that, with an HMD, participants were not correctly calibrated, in spatial terms, which resulted in many occurrences of collisions with the virtual environment. Adding an avatar of the self in the HMD view resulted in significantly fewer collisions. Collisions with the virtual environment did not occur in the CAVE, where the participants' own body was always present in the visual field (Lepecq, Bringoux, Pergandi, Coyle, & Mestre, 2009). This type of result suggests that participants need a colocalized representation of their own bodies to achieve precise spatial behavior, while wearing an HMD. There are many problems associated with this requirement, such as the need for precise biomechanical modeling and a realistic colocalized avatar.

The results from the above cited studies drove our choice to use a CAVE, in which participants naturally see their own bodies. The CAVE provides elements that are crucial to VR effectiveness, i.e. the sense of immersion and presence (Moore, Wiederhold, Wiederhold, & Riva, 2002). The sense of immersion, defined as “a sensorimotor coupling between a participant and a virtual environment” and presence, defined as “a psychological, attentional and cognitive state in which the participant, immersed within a virtual environment, behaves in accordance with the affordances provided by this environment […]”, are contingent on ecological validity and can lead to real-life behavior (Mestre, 2015, p.1). The sense of presence in the CAVE stems from the fact that not only are participants immersed visually and auditorily, but they can see their own bodies. Visually, they are surrounded by virtual images projected onto 3 or 4 screens (the floor and surrounding walls), providing a sensory illusion that creates a credible environment (for a review see Bohil, Alicea, & Biocca, 2011). In the CAVE, participants experience the sense of “agency” that arises from being able to gesture and move their arms and hands freely (Johnson-Glenberg, 2018), allowing them to perform more naturalistic and interactive tasks. The sensorimotor system is therefore much more fully engaged than in traditional experiments and elicited responses are closer to what probably occurs in real life (Bohil et al., 2011). Finally, the engaging aspect of this rich environment can also act as a motivational tool (Bayliss & Ballard, 2000).

For all of the above stated reasons, virtual reality is an attractive methodology to pair with EEG to study the interaction of motor and linguistic processing. Researchers have full control over multimodal sensory stimulation, making it possible to directly observe brain activity that correlates with specific types of sensory input, whether visual or auditory, in a more ecologically valid environment where naturalistic actions can be planned and performed. The intrinsic multimodality of human communication makes virtual reality paradigms particularly beneficial when studying language processing. In a recent review article of studies that used virtual reality in psycholinguistic research, Peeters (2019) claimed that what is most promising about virtual reality as an experimental tool is that it will “shift theoretical focus towards the interplay between different modalities in dynamic and communicative real-world environments, moving beyond and complementing studies that focus on one modality in isolation” (Peeters, 2019, p.6). Combining EEG with CAVE simulation of movement during linguistic processing can provide a novel and compelling view into how motor and linguistic systems may interact, which was the aim of the present work.

1.4. *The current study*

In the present study, we measured participants’ cortical activity while they listened to auditory action verbs and subsequently manipulated virtual objects or not in a CAVE. The main aim of this study was to provide evidence of μ ERSP during action language processing, in an ecological environment. In an effort to take into account theories of embodied cognition that argue that cognition is strongly constrained by one’s surrounding environment and physical state (Atkinson, 2010), we chose to use an ecologically realistic environment that required real and varied action. We used a CAVE which, compared to a real-word setting, allows for the controlled and synchronized presentation of stimuli with EEG, to the same degree as computerized experiments. This paradigm has the benefit of being more ecologically valid than traditional computerized set-ups. However, as outlined in a previous case report, recording EEG in a VR environment presents a particular technical challenge, due to both possible crosstalk between systems and participants’ movement (Török et al., 2014). Combining EEG recording and virtual stimulation also requires a precise synchronization process (Repetto, Cipresso et al., 2015). We were thus interested in providing a proof of concept, in addition to testing specific hypotheses about the role of motor activation during linguistic retrieval. To examine how motor processes may affect early linguistic processing, we explored the neural activity in the sensorimotor cortex during the auditory processing of verbs, prior to actual movement, via time-frequency analyses. We also examined ERP language related components. As concerns the overlap of linguistic and motor processes, we were specifically interested in ERD in the μ frequency bands (8–13 Hz) during the auditory processing of the verb. To address recent observations that μ ERD can be confounded with alpha ERD and that beta ERD provides a way to ensure that what is being observed is motor activation as opposed to alpha, we focused on both μ and beta ERD to show motor resonance (Hobson & Bishop, 2016). We hypothesized that we would observe μ ERD during verbal processing (Moreno et al., 2015); the distribution of the μ effect, whether central-parietal or more frontally located, as well as simultaneous beta ERD, should inform us of its nature (Hobson & Bishop, 2016).

Extant literature has revealed that activity in the 8–13 Hz frequency range is not a unitary measure (Klimesch, Doppelmayr, Pachinger, & Russegger, 1997) and can be divided into a lower range (8–10 Hz) and an upper range (11–13 Hz) and, in each range, ERD differs both in terms of its spatial distribution and in the processes thought to underlie it. ERD in the 8–10 Hz frequency band has been found to have a wide spatial distribution and to reflect processes related to attention and general task demands, which may thus be more reflective of alpha than of μ. In contrast, ERD in the 11–13 Hz frequency band has been revealed as being more topologically restricted and related to specific cognitive tasks, most notably processes related to semantic or long-term memory processes (Klimesch et al., 1997; Neuper & Pfurtscheller, 2001). We also examined the beta-band (13–30 Hz), which we subdivided into the following sub-bands: beta1 (13–18 Hz), beta2 (19–25 Hz) and beta3 (25–30 Hz). Previous research has revealed greater beta-band power suppression in response to action verbs compared to non-action verbs in the lower beta band, 13–25 Hz (Weiss, Berghoff, Rappelsberger, & Müller, 2001), such that sub-dividing the beta-band may allow us to disentangle beta activity related to language processing and that linked to motor activity.

We used a Go No-Go design in which participants either enacted the auditory verb upon subsequently presented virtual objects or not. This allowed us to examine sensorimotor activity during verbal processing (prior to movement) for both types of trials and to directly compare sensorimotor activity during verbal processing as a function of trial type. While we did not have a strong hypothesis concerning the effect of trial type, previous behavioral work has shown that manual responses are inhibited, on Go trials, if the go signal is presented simultaneously with a verbal stimulus denoting a hand movement (Sato et al., 2008). It is thus possible that greater μ ERD would be found in the present study for Go than No-Go trials if indeed motor preparation was inhibited by the semantic processing of the action verbs. The design also allowed us to determine whether variation in the μ ERD as a function of the type of trial (Go vs. No-Go) would be accompanied by a modification of ERP components, notably the contingent negative variation (CNV). Indeed there is debate concerning the direct coupling between these two responses (cf. Funderud et al., 2012; Filipović, Jahanshahi, & Rothwell, 2001; Zaepffel, Trachel, Kilavik, & Brochier, 2013).

2. Method

2.1. Participants

Twenty right-handed French native speakers (10 women, aged 20–26) participated in the study. Participants were volunteers from the student population of the Aix-Marseille Université, enrolled in the Science and Technique of Physical Sports Activities (STAPS) department. They had no history of neurological insult and received course credit in exchange for their participation. None had taken part in any prior VR experiment nor were they informed of the purpose of the experiment prior to the debriefing at the end of the session. All participants gave their written informed consent prior to the experiment, in keeping with the 1964 Helsinki Declaration, and the study was approved by the local ethics committee at Aix-Marseille Université.

2.2. Stimuli

Auditory stimuli consisted of 16 transitive French verbs (average number of phonemes = 5, ± 1) denoting actions that can be performed using one's hand and arm ("attraper" [catch], "cacher" [hide], "coucher" [lay down], "déplacer" [move], "empiler" [stack], "frotter" [rub], "lâcher" [let go of], "lancer" [throw], "faire pivoter" [pivot or twist], "pousser" [push], "relever" [make stand up], "secouer" [shake], "soulever" [pick up], "tapoter" [tap], "tirer" [pull], "faire tomber" [drop]). The auditory stimuli were produced by a trained female speaker and digitally recorded at 48 kHz (32-bit float) in a professional sound booth in a single session. They were subsequently spliced into individual tracks (Audacity software) and the duration of each auditory verb was determined. The verbs ranged in frequency per million from 1.16 to 415, with half being low frequency (average frequency = 17, ± 16) and the other half high frequency (average frequency = 211, ± 141). The choice of verbs was dictated both by their discriminability as concerns movement parameters and by the feasibility of tracking these movements with the finger-tracking glove. Visual stimuli consisted of 8 virtual geometric shapes (sphere, cube, cone, cylinder, rectangular prism, triangular prism, hexagonal prism, triangular pyramid). They were selected such that they did not provide affordances in relation to the set of verbs. The 16 auditory verbs were each presented 4 times (twice for each type of trial) in one of three pseudorandom orders. Each of the 8 objects was presented 16 times, equally often as a target and a distractor and across 10–12 verbs. The target object was color-coded green and the distractor was color-coded white.

2.3. Apparatus

2.3.1. The CAVE and the finger tracker

The Mediterranean Virtual Reality Center (CRVM) CAVE system consists of a cubic space measuring 3 × 3 × 3 × 4 m, with 3 vertical and 1 horizontal screen (floor). A graphic cluster of 4 video projectors delivers 4 stereoscopic projected images onto the 4 screens in real time (60 Hz). A movement capture system consisting of 8 infra red cameras provides the graphic cluster with the position of reflective targets on the user or on an entry peripheral device, allowing for interactions with the virtual environment. Participants wore 3D glasses and a 3-digit finger tracker (thumb, index and middle finger) on their right hand; the 3D glasses allowed participants to see their hand (and entire body) throughout the experiment and the finger tracker allowed for motion capture online as well as for participants to manipulate objects. The finger tracker was calibrated for each participant at the outset of the experiment to ensure the capture of acceptable movements, as predefined for each verb using UNITY. The apparatus used in this study is presented in Fig. 1.

Participants were visually surrounded by the virtual environment projected onto the four screens (the floor and 3 surrounding walls). The environment consisted of a virtual office containing a physical Plexiglas table on which the geometric objects to be manipulated were projected. Facing the participant was a large bookshelf containing objects typically found in offices such as books, filing boxes and framed pictures. To the left of the bookshelf was a virtual door and in the corner of the office was a virtual plant. Against the wall on the participant's left side was a chest of drawers bearing a vase and other office-type objects such as a filing folder. To the right of the participant was another chest of drawers.

2.3.2. Software

The UNITY software engine was used for stimulus presentation and a 64-channel Biosemi system (Actiview) was used for acquisition. The two systems were synchronized via a photodiode, which detected a change in luminosity (from black to white) of a square projected on the left bottom corner on the left vertical screen of the CAVE at the onset of each trial and again simultaneously to the onset of the virtual objects. The change in luminance was detected by the photodiode and the signal was sent to the acquisition system via one of the channels of the Biosemi AD system. The duration of the led signal was varied to distinguish the type of trial (Go vs. No-Go) and to indicate the side (left or right) of target object presentation.

Fig. 1. Participant in the CAVE, wearing a finger-tracker and 3D glasses while manipulating virtual objects.

2.3.3. EEG data acquisition

Electroencephalographic (EEG) activity was recorded continuously from 64 scalp electrodes located at left and right hemisphere positions over frontal, central, parietal, occipital, and temporal areas by means of a 64-channel electrode cap mounted with silver-chloride active electrodes (BioSemi Active Two system AD box). Individual electrodes were adjusted to a stable offset lower than 20 KΩ. Blinks and vertical eye movements (VEOG) were monitored via two external electrodes placed under each eye and horizontal eye movements (HEOG) were monitored via two electrodes positioned at the outer canthus of both eyes. External electrodes were placed over both the left and right mastoids and the left mastoid served as reference during EEG acquisition. EEG was sampled online at a rate of 2048 Hz; a band-pass filter (0.16–100 Hz) was applied online for visualization purposes only.

2.4. Procedure

Participants sat comfortably behind a Plexiglas desk, wearing the finger-tracking glove and 3D glasses. The session began with a 10-minute training phase during which participants learned how to manipulate the virtual objects. For this, they learned to use different hand positions: pinch using the index finger and the thumb to manipulate smaller objects, C-shaped hand to manipulate larger objects, flat open hand to tap, pull or push objects.

During the experimental phase, stimuli were presented in two blocks of 32 trials. At the beginning of each trial, the participant sat with his/her right hand in resting position, on top of a small textured circle placed in the center of the Plexiglas desk. A trial was initiated only when the participant's hand was detected in this position. A trial sequence began with the presentation of a visual prompt, projected onto the Plexiglas table, signaling the type of trial (✓ = Go; ✗ = No-Go), 500 msec prior to and throughout the auditory presentation of the verb (ex. "Tapote" [Tap]). Two and a half seconds after auditory verb onset, 2 different geometrical objects were projected onto the Plexiglas table: the target and distractor. On Go trials participants were told to perform the appropriate action on the target object. On No-Go trials they were instructed to simply listen to the verb. Post onset of the 2 objects, a constant 10-second period was allotted during which participants performed the action. A visual prompt was displayed for 2 s at the end of each trial during which participants were instructed to blink. The next trial was initiated when the participant's hand was detected in the resting position. If a movement was performed incorrectly on a Go trial or if the participant executed a Go during a No-Go trial, the trial was repeated at the end of the block. The experimental phase lasted roughly 20 min, with a short pause between the two blocks.

2.5. Data pre-processing

We used EEGLAB (Delorme & Makeig, 2004) to pre-process raw data. EEG data was downsampled to 512 Hz and bandpass filtered between 0.3 Hz and 80 Hz. The filtered data was re-referenced offline to the average of the two mastoids. Noisy electrodes were determined by calculating a robust noise adjusted z-score for each, as implemented in the ADJUST plugin for artifact detection (Mognon, Jovicich, Bruzzone, & Buiatti, 2011). This method calculates the ratio of the median absolute deviation of high frequency components (> 50 Hz) to low frequency components (< 50 Hz) for each electrode, expressed as a z-score relative to all other electrodes. Those electrodes with a robust z-score exceeding 5 were marked for possible rejection. This was complemented by visual examination of the power spectral density of each electrode to determine those with excessive low and high frequency activity or contaminated by line noise.

The continuous data was segmented into 3200 msec epochs, spanning 1200 msec before to 2000 msec after auditory verb onset. This trial length ensured sufficient data to resolve the low frequencies when carrying out time-frequency decomposition. However, for baseline correction a pre-stimulus interval of 200 msec (−250 msec to −50 msec) was applied and subsequent data analyses were limited to the 0 msec to 1000 msec post-stimulus interval.

Noisy electrodes marked for rejection were removed. Before carrying out independent component analysis (ICA), to correct for ocular movements, epochs that were highly contaminated with noise due, in particular, to movement were removed from the dataset. ICA was carried out on the segmented data of all 64 scalp electrodes for each participant using the infomax algorithm (Bell & Sejnowski, 1997) implemented in EEGLAB. Components corresponding to eye-blinks were determined automatically via the ADJUST Toolbox (Mognon et al., 2011) and, generally, only the first component was rejected. Epochs were then visually inspected again and those contaminated by noise were removed. A minimum of 53 (average of 57, ± 4) electrodes and 43 epochs (average of 52 ± 5) were retained per participant. At this point, rejected electrodes were interpolated using spherical spline interpolation. The data was then separated into Go and No-Go conditions. A total of 434 Go trials and 415 No-Go trails were retained overall, with an average of 25.5 ± 2.6 Go trials and 24.4 ± 4 No-Go trials per participant.

2.6. Event-Related Potential (ERP) analysis

For each participant, the mean over all trials was calculated to yield the subject-level ERP data. The ERP data of each subject was low-pass filtered, with a cutoff of 30 Hz. In line with previous electrophysiological studies of language processing (McLaughlin et al., 2010; Sneed-German, Herschensohn, & Frenck-Mestre, 2014) the 64 scalp electrodes were divided into 7 regions of interest (ROI): left frontal electrodes (AF3, F1, F3, F5), right frontal electrodes (AF4, F2, F4, F6), left frontal-central electrodes (FC1, FC3, FC5, C1, C3, C5), right frontal-central electrodes (FC2, FC4, FC6, C2, C4, C6), left central-parietal electrodes (CP1, CP3, CP5, P1, P3, P5), right central-parietal electrodes (CP2, CP4, CP6, P2, P4, P6), and midline electrodes (AFz, Fz, FCz, Cz, CPz, Pz). To determine the time-windows in which a significant difference between Go and No-Go trials emerged, a permutation test with false discovery rate (FDR) correction was carried out on all time points of the post-stimulus interval (0–1000 msec) for each electrode; to carry out the permutation, 1000 random partitions were performed. A significant difference was only taken into consideration ($q \leq 0.05$) if its duration exceeded 10 msec (8 consecutive time samples).

2.7. *Time-frequency decomposition*

The ERSP was computed for each participant both for Go and No-Go trials merged and for Go and No-Go trials separately using the Matlab toolbox, FieldTrip (Oostenveld, Fries, Maris, & Schoffelen, 2011). Time-frequency decomposition was carried out on a single trial basis as the squared norm of the convolution of the complex Morlet wavelet for each of the 64 scalp electrodes and for the 3–40 Hz frequency band. Within this frequency range, the wavelet width varied linearly from 3 to 10 cycles as a function of increasing frequency. This yielded a spectral bandwidth of 2 Hz and a temporal bandwidth of 318 msec at the lowest frequency of interest (3 Hz) and a spectral bandwidth of 8 Hz and a temporal bandwidth of 80 msec at the highest frequency of interest (40 Hz). At the single trial level, post-stimulus time-frequency data was z-score normalized relative to the pre-stimulus baseline interval (-250 to -50 msec). Then, for each participant, the grand average ERSP was calculated by averaging the trial-level ERSP and this data was entered into the subsequent statistical analyses.

2.8. *Time frequency statistical analyses*

2.8.1. *Post stimulus activity versus baseline*

To assess the statistical significance of ERSP in relation to the pre-stimulus baseline for Go and No-Go trials merged and for Go and No-Go trials separately, we applied a baseline permutation method (Delorme & Makeig, 2004) for each of 9 frontal central electrodes (*FC3, FC4, C3, C4, CP3, CP4, FCz, Cz and CPz*). These electrodes were chosen based on the hypothesis and findings of previous studies showing motor activation at these sites (Fargier et al., 2012; van Elk et al., 2010). This method involves permuting the pre-stimulus baseline values across both time and trials for each frequency to generate a surrogate distribution for each frequency value; we carried out 2000 permutations at each frequency. For each time-frequency point, values that fell within the 97.5% tail of the surrogate distribution were considered significant at $q \leq 0.05$. The comparison of Go and No-Go merged to baseline was computed across the 3–40 Hz range. In contrast, the independent comparison of Go to baseline and No-Go to baseline focused on the lower and upper μ bands, 8–10 Hz and 11–13 Hz, respectively.

2.8.2. *Cluster based permutation analyses: Go versus No-Go trials*

To directly compare Go and No-Go trials, the participant-level grand-average ERSP for each type of trial was entered into a cluster-based permutation analyses (Maris & Oostenveld, 2007). This non-parametric test simplifies the resolution of the multiple comparisons problem by correcting at the level of clusters that are determined based on an adjacency criterion and its calculation involved a multi-level statistical approach. At the first level, a dependent-samples *t*-test was performed for every data sample across conditions; data points corresponded either to time × frequency (for a given electrode) or electrode × time (for a given frequency band of interest) samples. A pre-set threshold of 5% (two-tailed) was used to group neighboring electrodes into clusters; neighbors were determined based on an adjacency criterion of a minimum of 2 electrodes calculated using the Delaunay triangulation function implemented in FieldTrip. To calculate cluster-level statistics, t-statistics were summed in each cluster and the maximum of the cluster-level statistic was determined. On the second level, we created a Monte-Carlo permutation distribution to calculate the significance probability. Participants' grand averages were randomly assigned to one of two conditions 1000 times and, for each random partition, the largest cluster-level statistic was determined. The Monte-Carlo permutation distribution was then constructed. The cluster-level test statistics were then compared to this permutation distribution and clusters in the highest or lowest 2.5th percentile of the distribution were considered significant (Fonteneau, Bozic, & Marslen-Wilson, 2015; Mazaheri et al., 2018).

Two comparisons were carried out. Cluster-based permutation analyses were carried out for two specific frequency bands of interest, the lower and upper μ bands (8–10 Hz and 11–13 Hz, respectively), for all 64 electrodes and the entire post-stimulus time interval (0–1000 msec), in which case each data sample constituted a spatial-temporal sample. The same analysis was carried out for individual electrodes of interest over the entire post-stimulus time interval and for the entire 3–40 Hz frequency band, in which case each data sample constituted a time × frequency sample.

3. Results

3.1. *Event-related potentials (ERPs)*

The mean ERPs over 6 ROIs as well as the 95% confidence intervals for both Go and No-Go conditions are presented in Fig. 2. A clear N1-P2 complex followed by an N400 can be seen, revealing clean recording of ERPs during the linguistic processing of action verbs in the CAVE. No statistically significant differences were revealed between the Go and No-Go conditions at any ROI or any time point.

3.2. *Event-related spectral perturbation (ERSP)*

3.2.1. *Go and No-Go merged*

Fig. 3 presents the ERSP of both trial types merged (Go + No-Go) for the 3–40 Hz frequency band and over the post-stimulus interval (0–1000 msec). We examined the mean oscillatory activity over groups of electrodes analyzed in previous studies that examined motor and semantic interactions (Fargier et al., 2012): Left frontal-central (FC1, FC3, FC5), Right frontal-central (FC2, FC4, FC6), Left Central (C1, C3, C5), Right Central (C2, C4, C6) and Left central-parietal (CP1, CP3, CP5) and Right central-parietal (CP2, CP4, CP6) and Midline electrodes (FCz, Cz, CPz).

The visual examination of the time-frequency maps presented in Fig. 3 revealed two patterns of activity. First, we see an early strong ERS emerging within the first 100 msec of the post-stimulus interval, which corresponds to the N1-P2 complex that we observed in our ERP results. In relation to this, it is important to underline that the time-frequency decomposition was carried out on a single-trial basis such that grand-average time-frequency activity includes both trial varying or induced activity (non-phase locked) and evoked activity (time locked but not phase locked in relation to stimulus onset) (Roach & Mathalon, 2008). Secondly, we observed an ERD in the μ band (8–13 Hz) that emerged in the 400–500 msec time window. Visual examination of this ERD revealed that it was stronger over the left hemisphere than the right, which is undoubtedly related to the fact that all participants were right-handed and executed actions with their right hand.

These patterns were substantiated by the statistical analysis of the post-stimulus interval (0–1000 msec) for all trials (Go + No-Go). Fig. 4 presents the results of this analysis; it reveals post-stimulus activity that was statistically significant ($q \leq 0.01$) compared to the baseline period. Only significant activity is shown. The μ-band ERD was statistically significant from the 400–500 msec time window up to 1000 msec after

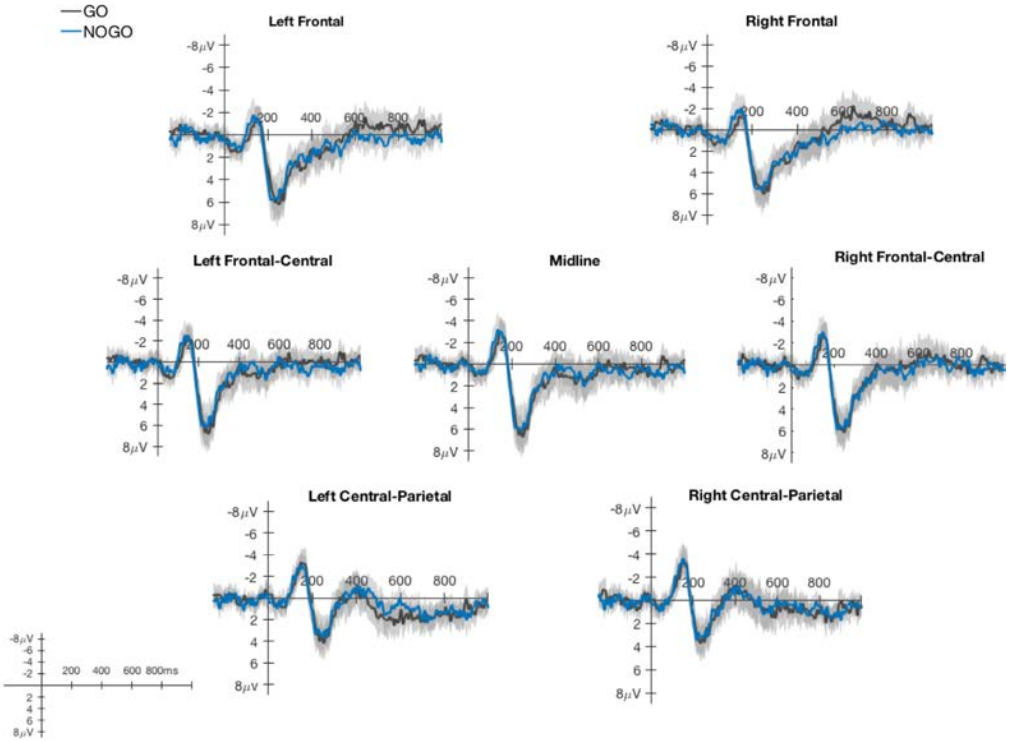


Fig. 2. Grand average ERPs and 95% confidence intervals for frontal electrodes (top), frontal-central electrodes (middle), and central-parietal electrodes (bottom), for Go and No-Go conditions.

stimulus onset. Results also showed significant ERD in the beta band, in particular in the 20–30 Hz frequency band, spanning beta 2 and 3.

3.2.2. *Go vs. No-Go comparison: permutation analyses*

To test the hypothesis of a difference in ERSP between Go and No-Go conditions, a two-tailed spatio-temporal cluster-based permutation test was carried out for each of the 9 electrodes of interest for the 1-second post-stimulus time window across the 3–40 Hz frequency band. Fig. 5 presents those 5 electrodes out of the 9 that revealed statistically significant differences. The time-frequency maps present the raw effect (Go/No-Go difference) and only statistically significant activity ($q \leq 0.025$) is shown. We found significantly greater ERD for Go vs. No-Go trials in the μ band from 700 to 1000 msec, for all electrodes except FC3 and C3 where it emerged earlier. In the 3 left hemisphere electrodes (FC3, C3 and CP3), the significant difference between conditions extended into the beta band (20–30 Hz, spanning beta 2 and beta 3) and, for FC3 and C3, emerged in an early time window (200–300 msec).

To test the hypothesis of a difference in ERD between conditions specifically in the μ band, a one-tailed spatio-temporal cluster-based permutation test was carried out for the low μ (8–10 Hz) and high μ (11–13 Hz) bands, over the entire post-stimulus time window (0–1000 msec) and over all 64 electrodes. The topographies in Fig. 6 present the log normalized Monte-Carlo significance probability or p-value ($-\log10(p)$) of the first significant cluster over consecutive 100 msec time windows for both μ bands. It is important to point out that, while the statistical analysis was carried out for each data sample ($\Delta t = 15.6$ msec), the results are presented in consecutive 100 msec time steps for visualization purposes. This reveals the spatial-temporal points for which the Go vs. No-Go difference was statistically significant ($q \leq 0.05$).

Greater ERD for Go trials compared to No-Go trials was revealed in both the 8–10 Hz ($p = .048$) and 11–13 Hz ($p = .05003$) frequency bands. For the 11–13 Hz frequency band, a significant difference emerged from 400 to 600 msec over a small number of central electrodes; the difference was also reliable from 800 msec to the end of the 1 sec time period. For the 8–10 Hz frequency band, the difference in μ ERD reached significance later, in the 600–700 msec time window over posterior electrodes. However, from the 700–800 msec time window to the end of the trial, this difference emerged over frontal-central electrodes.

3.2.3. *Significant ERSP for Go and No-Go independently*

In a final comparison, we calculated the significant ERSP for each condition independently, but focusing on significant ERD in the upper (11–13 Hz) and lower (8–10 Hz) μ bands. Results showed significant ERD in both Go and No-Go conditions and for both μ bands. Figs. 7a and 7b show significant ERSP post-stimulus activity for the low µ -band (8–10 Hz) and the upper μ band (11–13 Hz), respectively. Only significant ($q \leq 0.05$) post-stimulus activity is shown, all non-significant activity relative to the baseline is masked.

4. Discussion

4.1. *EEG-VR combination*

The present study examined the synergy between motor and semantic processes during language processing, in a novel protocol using EEG in an ecologically valid environment. Participants performed a Go/No-Go task in an interactive CAVE environment. They heard action verbs and subsequently saw virtual objects which they either manipulated or not. Our results showed, first, a clear pattern of language related ERPs during verb processing for all trials, i.e. an N1/P2 complex followed by an N400. As discussed in greater length below, we found no variation in the ERP response as a function of trial type (Go/No-Go).

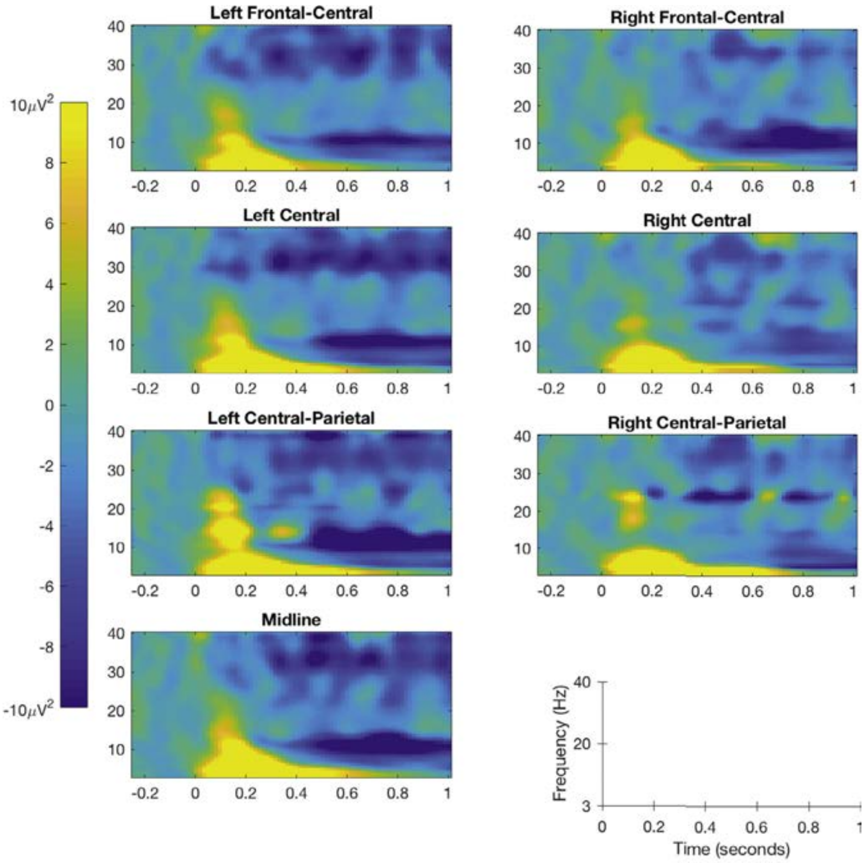


Fig. 3. Time-Frequency maps of all trial types (Go + No-Go).

Second, as detailed below, we found clear evidence of interactions between motor and linguistic processing, as shown by event-related desynchronization in language processing time windows. Significant ERD emerged during verb processing in both the μ (8–13 Hz) and beta band (20–30 Hz) for both Go and No-Go trials. While *μ ERD* emerged in the 400–500 msec time window, beta ERD emerged earlier (starting at 200 ms) and persisted. The direct comparison of Go to No-Go trials revealed greater μ ERD for Go trials. We did not find concomitant variation of the CNV and ERD. These results highlight the advantage of combining EEG and CAVE for the study of motor-language processes under well-controlled conditions and offer a novel, ecologically valid methodology to study these processes.

In comparison to the present study, it could be argued that previous EEG experiments that have focused on the neural signatures of motor-semantic interactions used relatively impoverished environments (Alemanno et al., 2012; Fargier et al., 2012; Moreno et al., 2015; von Nicolai et al., 2014; van Elk et al., 2010) and do not provide a clear reflection of how language is processed in real life (Knoeferle, 2015; Tromp et al., 2018). The present study overcame this limitation by presenting participants with a realistic albeit virtual environment in which they not only saw but actually manipulated virtual objects. In general, VR uses digital images and sound to create a credible and immersive sensory experience. Manual control tools such as finger-trackers allow participants to interact with objects and receive real-time feedback (Burdea & Coiffet, 2003) and head movements are tracked such that the visual environment responds to the participant's movement in a similar way to the real world. Compared to computer screen or head-mounted display systems, CAVE environments are highly effective in providing participants with a sense of presence and immersion (Juan & Pérez, 2009). In the current experiment, participants benefitted from an embodied experience as they were immersed in a virtual environment depicting an office, complemented by a physical Plexiglas table. They were equipped with a finger tracker, increasing the experience of object manipulation, although sensory feedback was not provided to avoid any confounds with the recording of motor activation. They were free to move their arms and hands and performed 16 naturalistic actions (ex. *throw, drop, push*) on virtual objects that obeyed natural physical constraints. Both the objects and the visual environment responded in real-time to participants' movements, providing them with a sense of agency (Johnson-Glenberg, 2018) and leading to real-life behavior (Mestre, 2015). Presence, immersion and agency all contribute to a greater involvement of the sensorimotor system, such that the responses elicited in these conditions are closer to what probably occurs in real life (Bohil et al., 2011). Combining the CAVE and EEG therefore enabled us to control multimodal sensory stimulation while observing the brain correlates of motor and linguistic interaction in an ecologically valid environment, where participants could plan and perform naturalistic movements. We did not, however, measure presence and thus have no concrete evidence that the present design afforded a greater sense of involvement than, say a flat screen presentation. Future studies would benefit from this added measure.

Combined EEG and VR can provide for a richer and more representative illustration of what occurs when motor and language processes overlap, and hence a better understanding of how language is

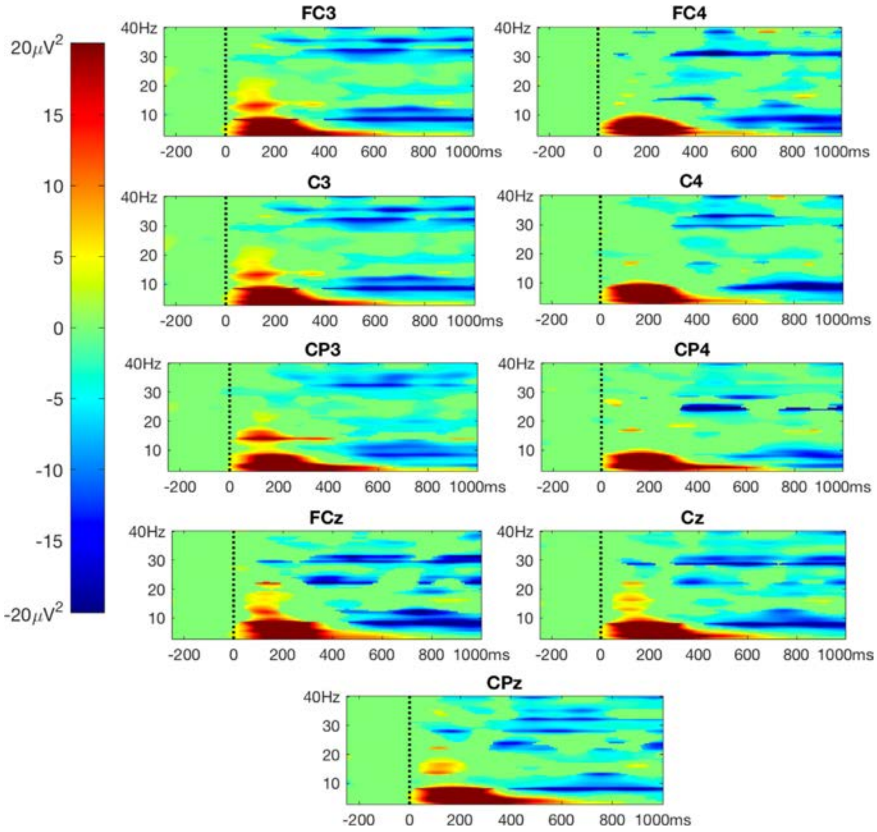


Fig. 4. Time-frequency maps showing statistically significant (q ≤ .05, FDR corrected) post-stimulus power relative to the baseline period for 9 electrodes over both hemispheres.

embodied. In relation to the technical constraints of the present work, it is noteworthy that we implemented a design in which participants remained stationary during the period of interest of EEG recording. This indeed afforded the extremely high quality of the ERP traces that were obtained in the present study. This choice in fact mirrors previous VR studies on language processing (Repetto, 2014; Repetto, Cipresso et al., 2015; Tromp et al., 2018) and is linked to the constraints of EEG as concerns the need to eliminate spurious noise. Nonetheless, in contrast to previous work, participants performed a variety of naturalistic movements. This thus afforded a far more interactive and ecological situation than previously used. One might nonetheless argue against the ecological validity of our experimental setup, because participants did not receive haptic feedback during the manipulation of virtual objects. This might create some cognitive dissonance for the participants and thereby influence the data. To assess this, a direct comparison with the manipulation of real objects would be required (cf. Repetto, Cipresso

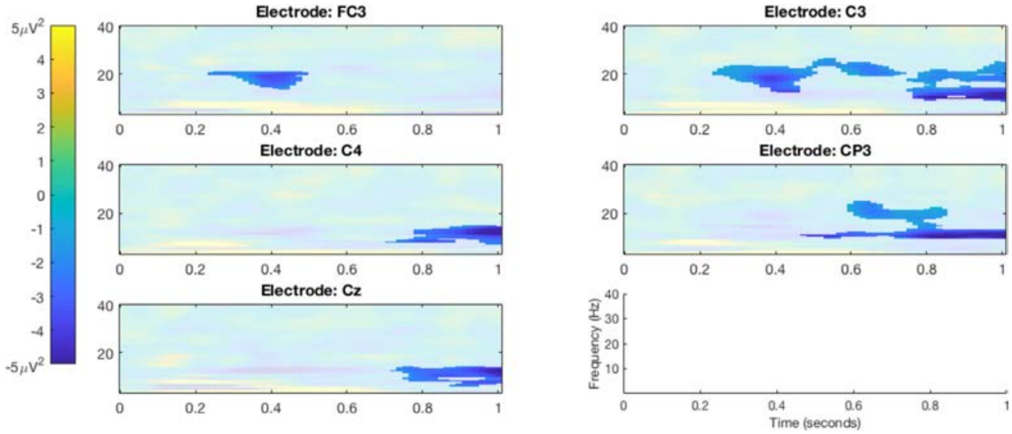


Fig. 5. Result of cluster-based permutation test comparing Go and No-Go trials across 3–40 Hz, showing the 5 electrodes that revealed statistically significant differences.

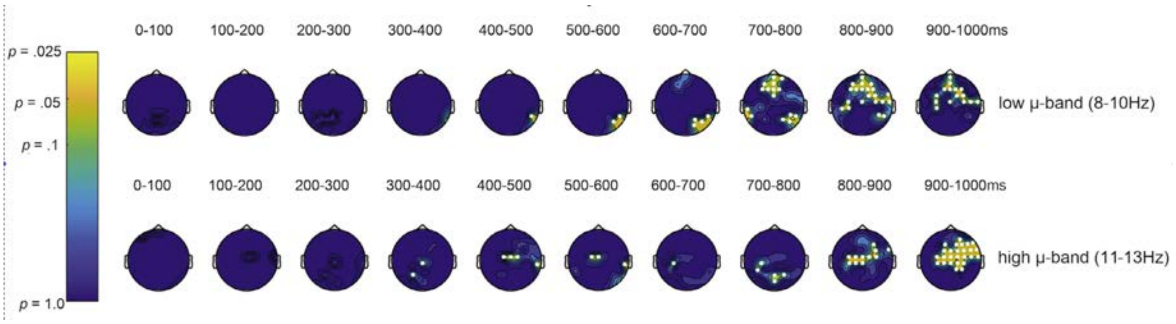


Fig. 6. Result of the spatio-temporal cluster-based permutation test comparing Go to No-trials over the entire post-stimulus time window and all 64 electrodes for the low (8–10 Hz) and high (11–13 Hz) *μ band.*

et al., 2015). However, our results refer to EEG recordings acquired *prior* to the onset of virtual objects and hence manipulation, which tends to counter the argument of cognitive dissonance. Nevertheless, Invitto, Faggiano, Sammarco, De Luca, and De Paolis (2016) report both behavioral and EEG results from a study in which they compared mental imagery, virtual manipulation and real grasping of objects. They found significant differences in ERP components (primarily N1) between real and virtual manipulation such that virtual action failed to facilitate perceptual processes compared to real action. These findings suggest a need for further investigation of the role of multimodal stimulation in deciphering the links between embodiment and linguistic processing. However, two important limitations of the comparison between our study and Invitto et al. (2016) are that they measured ERP in a “Go/No-Go” recognition task (respond only if the stimulus had been presented) and *after* the manipulation (training) phase. Hence, their “Go/No-Go” task was not similar to ours and undoubtedly measured different aspects of processing.

4.2. Interaction between motor and semantic processes

The current study explored motor activation in response to auditory action verbs as well as the effect of motor planning on linguistic processing. This was investigated in part by modulations in the ERP signature. Previous behavioral (Buccino et al., 2005; Glenberg & Kaschak, 2002; Sato et al., 2008) and EEG (Aravena et al., 2010) studies have shown that processing action verbs is facilitated when congruent actions are planned prior to verb presentation but hindered when congruent actions are planned simultaneously. We therefore expected No-Go trials to elicit a greater N400 compared to Go trials, possibly due to interference effects linked to inhibiting action (García & Ibáñez, 2016). In line with Aravena et al. (2010), we posited that Go trials might also increase early, motor-related ERPs, showing enhanced motor preparation. Contrary to these predictions, the analyses of ERPs revealed no differences in motor preparation (RP) or semantic processing (N400) across conditions. One possible explanation for the absence of differences in motor preparation is that, whereas in Aravena et al.’s experiment (2010) participants planned a specific and repeated action (closed

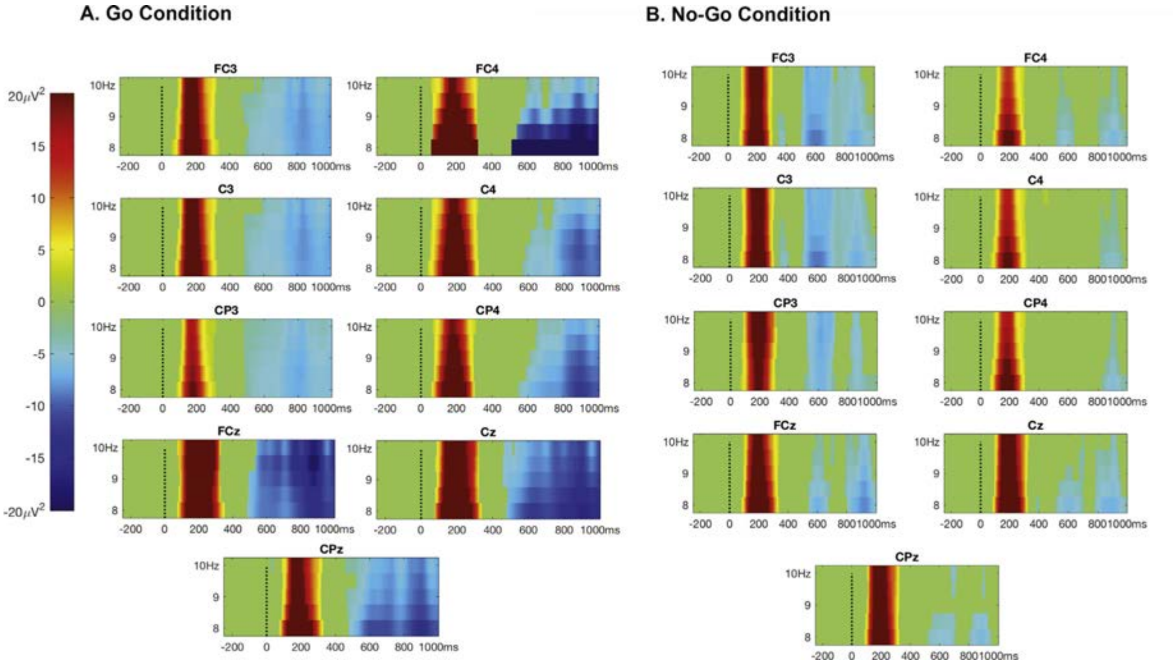


Fig. 7a. Post-stimulus oscillatory activity over the 8–10 Hz frequency range for (A) Go trials and (B) No-Go trials. Only statistically significant ($q \le 0.05$, FDR corrected) ERSP values are shown.

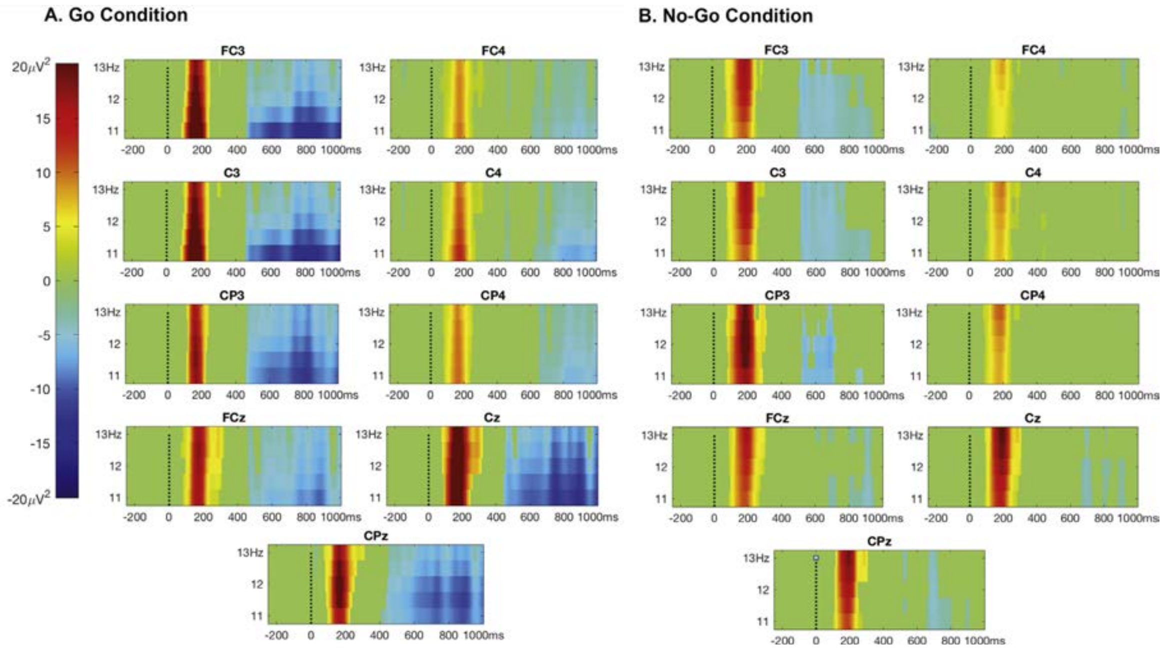


Fig. 7b. Post-stimulus oscillatory activity over the 11–13 Hz frequency range for (A) Go trials and (B) No-Go trials. Only statistically significant ($q \leq 0.05$, FDR corrected) ERSP values are shown.

or open-handed manual response throughout the experiment), in our experiment participants were requested to produce one of numerous hand movements, which they had to plan during a 2.5 s period before object presentation. It is possible that planning to perform a pre-specified repetitive movement allowed participants to pre-program movements in a way that was not possible under the current conditions, in which participants performed specific but variable actions on Go trials. The lack of an N400 effect is open to speculation; however, in the current design all actions were coherent, the only variation being the specific hand movement executed for a given verb. This differs from previous work in which participants knew a priori which movement to perform across all trials and had to semantically integrate the specific verb with a specific action on a trial to trial basis (Aravena et al., 2010). In the same line of argumentation, it is possible that the facilitation of semantic integration by coherent motor preparation only emerges under optimal conditions in which participants know a priori the movement to execute. It is noteworthy that the finding of facilitation in ACE studies is indeed not systematic but subject to numerous constraints, including the timing of motor preparation (Sato et al., 2008), the linguistic context (Aravena et al., 2012; Boulenger et al., 2006) and even the syntactic formulation (Aravena et al., 2014).

4.3. Modulation of ERSP

ERPs do not tell the whole story when it comes to motor activation as they use phase-locked linear averaging methods and hence do not capture the large portion of motor cortex signal that is not phase-locked to the stimulus (Pfurtscheller & Lopes da Silva, 1999; Vukovic & Shtyrov, 2014). Event-related ERS/ERD, however, considers both phase-locked and unlocked cortical activity. In the present study we capitalized on ERSP to explore neural activity in the sensorimotor cortex, notably as reflected by ERD in the μ frequency bands (8–13 Hz) in response to action verb processing and prior to physical movement. Our results revealed greater motor-related cortical activity (μ ERD) during verb processing for both Go and No-Go trials, starting in the 400–500 msec time window.

Previous studies have provided evidence of early activation of motor areas during linguistic processing. Pulvermüller, Härle, and Hummel (2001) used high-density EEG to measure cortical activity while participants read action verbs performed using different body parts. Somato-specific activation was found along the motor strip starting at 250 msec after verb presentation. In line with these findings, one could hypothesize that ERD in the μ frequency bands during verb processing should have occurred in our experiment in this window. Unlike the Pulvermüller et al. (2001) study, which used written single verbs as stimuli, we used auditory stimuli. The timing of lexical access is notoriously more difficult to determine for auditory language compared to written language (Hauk, Shtyrov, & Pulvermüller, 2008); this could account for the differences across studies.

The comparison of Go and No-Go trials to baseline revealed significant μ ERD starting in the 400-500 msec time window, which is generally considered to be associated with lexical access (Indefrey & Levelt, 2004; Kutas & Federmeier, 2011). This result is in line with that reported by Fargier et al. (2012) in a word learning paradigm, where no ERD was observed in the μ band for novel words prior to training but emerged following two training sessions in which the novel words were associated with hand movements, and was observed for centro-parietal electrodes starting at 450 msec post stimulus onset. It is important to note that Fargier et al. (2012) included a second day of training and test sessions for which the results were less conclusive. Our study looked at motor activation during the processing of well-known words (verbs) in the native language, such that we would not expect the pattern of results to vary over time. We also found significant ERD in the beta band, in particular in the 20–30 Hz band (spanning beta 2 and 3) starting from 200 msec for frontal central electrodes. As discussed below, the co-occurrence of ERD for the μ and beta bands provides an argument against the hypothesis that the μ ERD we report was in fact alpha. The fact that motor resonance was found in both conditions when they were tested independently also suggests that motor activation was not (only) caused by action planning but, in line with previous studies, was involved in semantic processing (Aravena et al., 2010; Boulenger et al., 2006; Pulvermüller et al., 2001).

We also investigated differences in motor activation during verb processing as a function of Trial type (Go vs. No-Go) for both the lower (8–10 Hz) and higher μ band (11–13 Hz). The lower μ band has been associated with action observation and the higher μ band with action execution (Aridan, Ossmy, Buaron, Reznik, & Mukamel, 2018). Analyses revealed greater μ ERD for Go compared to No-Go trials across the two frequency bands but for different time windows and with different distributions. In the higher μ band (11–13 Hz), greater ERD for Go trials emerged in the 400–500 msec time window for central-parietal electrodes. Although actual movement was delayed to 2500 msec after verb onset during Go trials, the go signal was presented only 200 msec prior to verb presentation. Previous Go/No-Go studies have shown that when the go signal was presented simultaneously with a verb describing a motion with a specific effector, responses using that same effector were slower (Buccino et al., 2005; Sato et al., 2008). In the present experiment, it is possible that Go trials produced greater μ ERD because processing action verbs inhibited motor preparation. In the lower μ band (8–10 Hz), a greater ERD for Go trials emerged later, starting at 700 msec, and was concentrated primarily over central and frontal electrodes. This activity is likely not a direct reflection of motor activation but indicative of activation in “convergence zones” of language and motor structures (Damasio, 1989; Fargier et al., 2012).

Finally, it is of interest to note that the significant variations we observed in ERSP were not accompanied by significant modulations in the contingent negative variation (CNV). Various studies have examined the relationship between the reduction in spectral power in the alpha and beta bands (ERD) and the increase CNV in Go/No-Go paradigms (Filipović et al., 2001; Funderud et al., 2012; Mento, 2013; Zaepffel et al., 2013). The CNV is typically seen in paradigms where participants receive a warning signal (S1) followed by a target stimulus (S2) and is thought to reflect a series of intentional motor, preparatory and decisional processes (Funderud et al., 2012; Mento, 2013). It is often separated into an earlier and a later component. The late (or terminal) CNV, beginning up to 1.5 s before S2, causes activity over frontal and prefrontal cortices and has been associated with sustained alpha (8–13 Hz) and beta (14–30 Hz) ERD/ERS (Morash, Bai, Furlani, Lin, & Hallett, 2008). The association of increased CNV amplitudes and reduced event-related spectral power in the alpha, beta, theta and low gamma bands for Go versus the No-Go trials has been interpreted as possibly illustrating a coordinated, dynamic change in neural networks involved in motor preparation (Funderud et al., 2012). However, a number of studies have provided evidence against a direct coupling between these two phenomena by showing discrepancies between them and arguing that they reflect different cognitive and motor processes (Filipović et al., 2001; Zaepffel et al., 2013). In our experiment, greater μ synchronization was found for Go versus No-Go trials but no significant differences between conditions emerged for the CNV. This could be taken as evidence that the ERD and CNV reflect different cognitive and motor processes. However, our study did show some variation in the CNV albeit very small, which may have been too slight to detect with the sample size of the current study (a G-power test based on our sample size and Cohen’s d suggested a population of over 400 would be needed to produce a significant result). As such, our study cannot adjudicate this question.

4.4. Caveats

One important limitation of the present study is the lack of a control condition using abstract verbs. Such a condition would have allowed us to see whether motor resonance emanated from processing action verbs specifically. Based on previous studies, we posit that we would have found greater μ rhythm ERD for action verbs compared to abstract verbs (Alemanno et al., 2012; Moreno et al., 2015). Importantly, a non-action verb control condition would also have allowed us to observe motor activation during Go vs. No-Go trials for action versus non-action verbs, once again illustrating whether the greater motor activation we found for Go compared to No-Go trials originated from interference from the action semantic content of the verbs as opposed to motor preparation alone. Note that Sato et al. (2008) based their conclusions of semantic interference on motor processing on the comparison of verbs related to a specific effector (hand vs. foot), not on action verbs compared to abstract verbs, despite having included abstract controls (but see Buccino et al., 2005). This could also provide an avenue to explore, i.e. the inclusion of movements for different effectors (cf. Buccino et al., 2005).

Another caveat to bear in mind is the current debate on what μ ERD reveals (Hobson & Bishop, 2016). Despite the growing use of time-locked μ ERD as a marker of motor neuron activity (Moreno et al., 2015), μ and alpha frequency bands (8–13 Hz) overlap and it is therefore important to dissociate the two (Hobson & Bishop, 2016). One way to accomplish this is to focus on distribution. Mu band activity is thought to originate in the motor and premotor cortex, measured in centro-parietal sites. Alpha band activity, on the other hand, is reportedly found over occipital and frontal sites (Fargier et al., 2012; Moreno et al., 2015). In addition, as beta band activity is thought to reflect motor activation directly (Pfurtscheller & Lopes da Silva, 1999), ERD in the beta band generally accompanies μ ERD and has often been considered as an indicator that what is being detected is indeed μ and not alpha ERD (van Elk et al., 2010). We found greater μ ERD compared to baseline for both Go and No-Go trials as well as both beta 2 (19–25 Hz) and beta 3 (25–30 Hz) bands compared to baseline. This, along with the finding that ERD in the μ band was greater for central compared to posterior electrodes, indicates that the effects are indeed linked to μ and were not confounded with posterior alpha.

5. Conclusion

The present study used a naturalistic setting to investigate motor activation during language processing within an embodied framework. Results showing motor activation in time windows associated with semantic processing are in line with the language studies that found μ-band ERD compared to baseline during lexical-semantic retrieval of action language (Alemanno et al., 2012; Fargier et al., 2012; Moreno et al., 2015; von Nicolai et al., 2014; van Elk et al., 2010). The greater action-related μ ERD during verb processing for all trials, but prior to movement proper, also bolsters the claim that sensorimotor processing is involved in the conceptual representation of linguistic information. Our results indicating greater μ ERD related to single verb processing for Go compared to No-Go trials are in line with ACE studies showing that movement preparation interacts with semantic processing (Aravena et al., 2010; Buccino et al., 2005; Sato et al., 2008). The present results, which used auditory verbs in conjunction with the execution of the specified action to measure the overlap of motor and linguistic processing, bear strong similarity to those reported by Fargier et al. (2012) who also used a varied set of linguistic materials and associated actions to examine this question. However, in the present study participants engaged with virtual objects rather than observed movements. Our results further validate the use of time-frequency analysis to measure motor activation in this novel EEG-CAVE experimental paradigm.

Acknowledgement

This work, carried out within the Labex BLRI (ANR-11-LABX-0036) and the Institut Convergence ILCB (ANR-16-CONV-0002), has benefited from support from the French government, managed by the French National Agency for Research (ANR) and the Excellence Initiative of Aix-Marseille University (A*MIDEX).

References

Alemanno, F., Houdayer, E., Cursi, M., Velikova, S., Tettamanti, M., Comi, G., & Leocani,

L. (2012). Action-related semantic content and negation polarity modulate motor areas during sentence reading: An event-related desynchronization study. *Brain Research, 1484*, 39–49.

Aravena, P., Courson, M., Frak, V., Cheylus, A., Paulignan, Y., Deprez, V., & Nazir, T. (2014). Action relevance in linguistic context drives word-induced motor activity. *Frontiers in Human Neuroscience, 8*, 163.

Aravena, P., Delevoye-Turrell, Y., Deprez, V., Cheylus, A., Paulignan, Y., Frak, V., & Nazir, T. (2012). Grip force reveals the context sensitivity of language-induced motor activity during "action words" processing: Evidence from sentential negation. *PLoS One, 7*(12), e50287.

Aravena, P., Hurtado, E., Riveros, R., Cardona, J. F., Manes, F., & Ibáñez, A. (2010). Applauding with closed hands: Neural signature of action-sentence compatibility effects. *PloS One, 5*(7), e11751.

Aridan, N., Ossmy, O., Buaron, B., Reznik, D., & Mukamel, R. (2018). Suppression of EEG mu rhythm during action observation corresponds with subsequent changes in behavior. *Brain Research, 1691*, 55–63.

Atkinson, D. (2010). Extended, embodied cognition and second language acquisition. *Applied Linguistics, 31*(5), 599–622.

Barsalou, L. W. (1999). Perceptions of perceptual symbols. *Behavioral and Brain Sciences, 22*(4), 637–660.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*(1), 617–645.

Bayliss, J. D., & Ballard, D. H. (2000). A virtual reality testbed for brain-computer interface research. *IEEE Transactions on Rehabilitation Engineering, 8*(2), 188–190.

Bell, A. J., & Sejnowski, T. J. (1997). The "independent components" of natural scenes are edge filters. *Vision Research, 37*(23), 3327–3338.

Bohil, C. J., Alicea, B., & Biocca, F. A. (2011). Virtual reality in neuroscience research and therapy. *Nature Reviews. Neuroscience, 12*(12), 752–762.

Boulenger, V., Roy, A. C., Paulignan, Y., Deprez, V., Jeannerod, M., & 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.

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

Burdea, G. C., & Coiffet, P. (2003). *Virtual reality technology.* John Wiley & Sons.

Caetano, G., Jousmäki, V., & Hari, R. (2007). Actor's and observer's primary motor cortices stabilize similarly after seen or heard motor actions. *Proceedings of the National Academy of Sciences of the United States of America, 104*(21), 9058–9062.

Damasio, A. R. (1989). Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition, 33*(1–2), 25–62.

de Vega, M., Moreno, V., & Castillo, D. (2013). The comprehension of action-related sentences may cause interference rather than facilitation on matching actions. *Psychological Research, 77*(1), 20–30.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134*(1), 9–21.

Diefenbach, C., Rieger, M., Massen, C., & Prinz, W. (2013). Action-sentence compatibility: The role of action effects and timing. *Frontiers in Psychology, 4*, 272.

Fargier, R., Paulignan, Y., Boulenger, V., Monaghan, P., Reboul, A., & Nazir, T. A. (2012). Learning to associate novel words with motor actions: Language-induced motor activity following short training. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior, 48*(7), 888–899.

Filipović, S. R., Jahanshahi, M., & Rothwell, J. C. (2001). Uncoupling of contingent negative variation and alpha band event-related desynchronization in a go/no-go task. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology, 112*(7), 1307–1315.

Fodor, J. A. (1980). Methodological solipsism considered as a research strategy in cognitive psychology1. *Behavioral and Brain Sciences, 3*(1), 63–73.

Fodor, J. A. (1987). *Psychosemantics: The Problem of Meaning in the Philosophy of Mind.* MIT Press.

Fonteneau, E., Bozic, M., & Marslen-Wilson, W. D. (2015). Brain network connectivity during language comprehension: Interacting linguistic and perceptual subsystems. *Cerebral Cortex (New York, N.Y.: 1991), 25*(10), 3962–3976.

Funderud, I., Lindgren, M., Løvstad, M., Endestad, T., Voytek, B., Knight, R. T., & Solbakk, A.-K. (2012). Differential Go/NoGo activity in both contingent negative variation and spectral power. *PLoS ONE, 7*(10).

García, A. M., & Ibáñez, A. (2016). A touch with words: Dynamic synergies between manual actions and language. *Neuroscience and Biobehavioral Reviews, 68*, 59–95.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review, 9*(3), 558–565.

Goldin-Meadow, S. (2011). Learning through gesture. *Wiley Interdisciplinary Reviews: Cognitive Science, 2*(6), 595–607.

Grandchamp, R., & Delorme, A. (2011). Single-trial normalization for event-related spectral decomposition reduces sensitivity to noisy trials. *Frontiers in Psychology, 2*, 236.

Graziano, M., & Gullberg, M. (2018). When speech stops, gesture stops: Evidence from developmental and crosslinguistic comparisons. *Frontiers in Psychology, 9*.

Hauk, O., Shtyrov, Y., & Pulvermüller, F. (2008). The time course of action and action-word comprehension in the human brain as revealed by neurophysiology. *Journal of Physiology-Paris, 102*(1), 50–58.

Hobson, H. M., & Bishop, D. V. M. (2016). Mu suppression – A good measure of the human mirror neuron system? *Cortex, 82*, 290–310.

Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition, 92*(1â€"2), 101–144.

Invitto, S., Faggiano, C., Sammarco, S., De Luca, V., & De Paolis, L. T. (2016). Haptic, virtual interaction and motor imagery: Entertainment tools and psychophysiological testing. *Sensors (Basel, Switzerland), 16*(3).

Johnson-Glenberg, M. C. (2018). Immersive VR and education: Embodied design principles that include gesture and hand controls. *Frontiers in Robotics and AI, 5*.

Juan, M. C., & Pérez, D. (2009). Comparison of the levels of presence and anxiety in an acrophobic environment viewed via HMD or CAVE. *Presence: Teleoperators and Virtual Environments, 18*(3), 232–248.

Kaschak, M. P., & Borreggine, K. L. (2008). Temporal dynamics of the action–sentence compatibility effect. *Quarterly Journal of Experimental Psychology, 61*(6), 883–895.

Klimesch, W., Doppelmayr, M., Pachinger, T., & Russegger, H. (1997). Event-related desynchronization in the alpha band and the processing of semantic information. *Brain Research. Cognitive Brain Research, 6*(2), 83–94.

Knoeferle, P. (2015). Language comprehension in rich non-linguistic contexts: Combining eye-tracking and event-related brain potentials. In R. M. Willems (Ed.). *Cognitive neuroscience of natural language use* (pp. 77–100). Cambridge: Cambridge University Press.

Koelewijn, T., van Schie, H. T., Bekkering, H., Oostenveld, R., & Jensen, O. (2008). Motor-cortical beta oscillations are modulated by correctness of observed action. *NeuroImage, 40*(2), 767–775.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62*, 621–647.

Ladouce, S., Donaldson, D. I., Dudchenko, P. A., & Ietswaart, M. (2017). Understanding minds in real-world environments: Toward a mobile cognition approach. *Frontiers in Human Neuroscience, 10*.

Legault, J., Zhao, J., Chi, Y.-A., Chen, W., Klippel, A., & Li, P. (2019). Immersive virtual reality as an effective tool for second language vocabulary learning. *Languages, 4*(1), 13.

Lepecq, J.-C., Bringoux, L., Pergandi, J.-M., Coyle, T., & Mestre, D. (2009). Afforded actions as a behavioral assessment of physical presence in virtual environments. *Virtual Reality, 13*(3), 141–151.

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology, Paris, 102*(1–3), 59–70.

Makeig, S. (1993). Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones. *Electroencephalography and Clinical Neurophysiology, 86*(4), 283–293.

Makeig, S., Debener, S., Onton, J., & Delorme, A. (2004). Mining event-related brain dynamics. *Trends in Cognitive Sciences, 8*(5), 204–210.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164*(1), 177–190.

Matsumoto, J., Fujiwara, T., Takahashi, O., Liu, M., Kimura, A., & Ushiba, J. (2010). Modulation of mu rhythm desynchronization during motor imagery by transcranial direct current stimulation. *Journal of Neuroengineering and Rehabilitation, 7*, 27.

Mazaheri, A., Segaert, K., Olichney, J., Yang, J.-C., Niu, Y.-Q., Shapiro, K., & Bowman, H. (2018). EEG oscillations during word processing predict MCI conversion to Alzheimer's disease. *NeuroImage: Clinical, 17*, 188–197.

McLaughlin, J., Tanner, D., Pitkänen, I., Frenck-Mestre, C., Inoue, K., Valentine, G., & Osterhout, L. (2010). Brain potentials reveal discrete stages of L2 grammatical learning. *Language Learning, 60*(Suppl. 2), 123–150.

Mento, G. (2013). The passive CNV: Carving out the contribution of task-related processes to expectancy. *Frontiers in Human Neuroscience, 7*.

Mestre, D. R. (2015). On the usefulness of the concept of presence in virtual reality applications. *The engineering reality of virtual reality 2015* (pp. 93920J). International Society for Optics and Photonics.

Mestre, D. R., Louison, C., & Ferlay, F. (2016). *The contribution of a virtual self and vibrotactile feedback to walking through virtual apertures. Proceedings of HCII, Toronto, 17–22 July*.

Mognon, A., Jovicich, J., Bruzzone, L., & Buiatti, M. (2011). ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology, 48*(2), 229–240.

Moore, K., Wiederhold, B. K., Wiederhold, M. D., & Riva, G. (2002). Panic and agoraphobia in a virtual world. *Cyberpsychology & Behavior: The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society, 5*(3), 197–202.

Morash, V., Bai, O., Furlani, S., Lin, P., & Hallett, M. (2008). Classifying EEG signals preceding right hand, left hand, tongue, and right foot movements and motor imageries. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology, 119*(11), 2570–2578.

Moreno, I., de Vega, M., León, I., Bastiaansen, M., Glen Lewis, A., & Magyari, L. (2015). Brain dynamics in the comprehension of action-related language. A time-frequency analysis of mu rhythms. *NeuroImage, 109*, 50–62.

Neuper, C., & Pfurtscheller, G. (2001). Event-related dynamics of cortical rhythms: Frequency-specific features and functional correlates. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology, 43*(1), 41–58.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience, 2011*, 156869.

Peeters, D. (2019). Virtual reality: A game-changing method for the language sciences. *Psychonomic Bulletin & Review, 26*(1), 1–7.

Pfurtscheller, G., & Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: Basic principles. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology, 110*(11), 1842–1857.

Pineda, J. A. (2005). The functional significance of mu rhythms: Translating "seeing" and "hearing" into "doing". *Brain Research. Brain Research Reviews, 50*(1), 57–68.

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews. Neuroscience, 6*(7), 576–582.

Pulvermüller, F., Härle, M., & Hummel, F. (2001). Walking or talking? Behavioral and neurophysiological correlates of action verb processing. *Brain and Language, 78*(2), 143–168.

Repetto, C. (2014). The use of virtual reality for language investigation and learning. *Frontiers in Psychology, 5*.

Repetto, C., Cipresso, P., & Riva, G. (2015). Virtual action and real action have different impacts on comprehension of concrete verbs. *Frontiers in Psychology, 6*.

Repetto, C., Colombo, B., & Riva, G. (2015). *Is motor simulation involved during foreign language learning? A virtual reality experiment.* SAGE Open.

Roach, B. J., & Mathalon, D. H. (2008). Event-related EEG time-frequency analysis: An overview of measures and an analysis of early gamma band phase locking in schizophrenia. *Schizophrenia Bulletin, 34*(5), 907–926.

Sato, M., Mengarelli, M., Riggio, L., Gallese, V., & Buccino, G. (2008). Task related modulation of the motor system during language processing. *Brain and Language, 105*(2), 83–90.

Sneed-German, E., Herschensohn, J., & Frenck-Mestre, C. (2014). *Pronoun processing in Anglophone late L2 learners of French: Behavioral and ERP evidence. EuroSLA 20. Presented at the York, United Kingdom*.

Spadacenta, S., Gallese, V., Fragola, M., & Mirabella, G. (2014). Modulation of arm reaching movements during processing of arm/hand-related action verbs with and without emotional connotation. *PLOS ONE, 9*(8), e104349.

Toni, I., de Lange, F. P., Noordzij, M. L., & Hagoort, P. (2008). Language beyond action. *Journal of Physiology-Paris, 102*(1), 71–79.

Török, Á., Sulykos, I., Kecskés-Kovács, K., Persa, G., Galambos, P., Kóbor, A., & Honbolygó, F. (2014). Comparison between wireless and wired EEG recordings in a virtual reality lab: Case report. *2014 5th IEEE conference on cognitive infocommunications (CogInfoCom)* (pp. 599–603). .

Tromp, J., Peeters, D., Meyer, A. S., & Hagoort, P. (2018). The combined use of virtual reality and EEG to study language processing in naturalistic environments. *Behavior Research Methods, 50*(2), 862–869.

van Elk, M., van Schie, H. T., Zwaan, R. A., & Bekkering, H. (2010). The functional role of motor activation in language processing: Motor cortical oscillations support lexical-semantic retrieval. *NeuroImage, 50*(2), 665–677.

von Nicolai, C., Engler, G., Sharott, A., Engel, A. K., Moll, C. K., & Siegel, M. (2014). Corticostriatal coordination through coherent phase-amplitude coupling. *Journal of Neuroscience, 34*(17), 5938–5948.

Vukovic, N., & Shtyrov, Y. (2014). Cortical motor systems are involved in second-language comprehension: Evidence from rapid mu-rhythm desynchronisation. *NeuroImage, 102*(Pt 2), 695–703.

Weiss, S., Berghoff, C., Rappelsberger, P., & Müller, H. M. (2001). Elektrophysiologische Hinweise zur Kategorisierung von Verben. *Tagungsband der 1. Jahrestagung der Gesellschaft für Aphasieforschung und -behandlung GAB* (pp. S.64). .

Wilson, A. D., & Golonka, S. (2013). Embodied cognition is not what you think it is. *Frontiers in Psychology, 4*, 58.

Zaepffel, M., Trachel, R., Kilavik, B., & Brochier, T. (2013). Modulations of EEG beta power during planning and execution of grasping movements. *PLoS ONE, 27* 11877–11877.

Chapter 5

Brain Waves in Virtual Reality

Bolger, D., Pergandi, J.-M., Zappa, A., Mallet, P., Dubarry, A.-S, Frenck-Mestre, C., & Mestre, D., submitted to *Conference on Human Factors in Computing Systems*

The present paper describes a novel experimental protocol that combines electroencephalography (EEG) and an interactive 3-D immersive environment (CAVE). The specific goal was to study the neural interactions between linguistic and motor processing, but the applications of the design are numerous. The CAVE provides greater ecological validity than traditional laboratory settings while still allowing for the control of experimental stimuli. In our study, EEG was recorded while participants processed auditory linguistic stimuli and manipulated virtual objects while wearing a finger-tracking device and 3D glasses in a 4-projector screen CAVE. The obtained EEG data showed clean, clear language-related ERP (event-related potential) components as well as modulations of the EEG signal as a function of experimental manipulations (Anonymous). We provide a detailed description of the design, technical set-up and how EEG (Biosemi system) and the CAVE setup (running Unity and custom software), were successfully synchronized.

5.1 Introduction

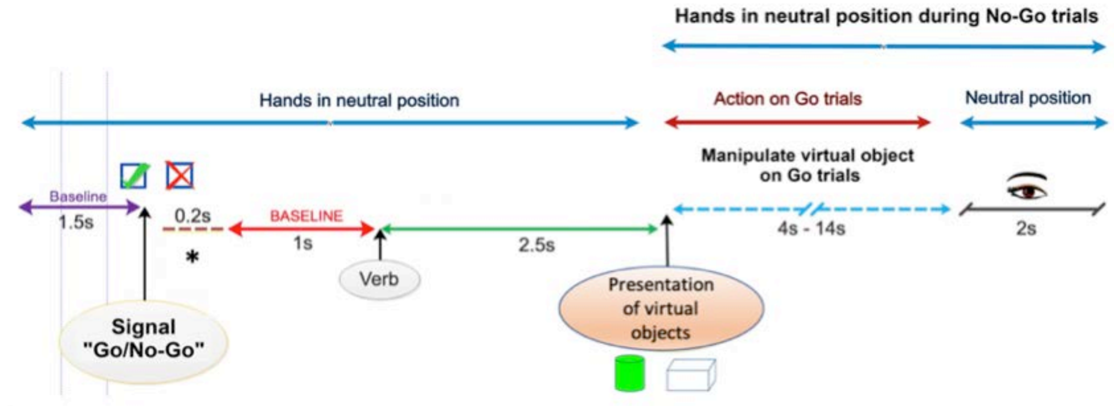
In search of greater ecological validity in laboratory experiments, Virtual Reality (VR) has long been proposed as a means to reduce the distance between the need to control experimental factors and real-life situations (Loomis, Blascovich & Beall, 1999; Tarr & Warren, 2002), enabling participants to act and feel like they actually are in the virtual environment (Sanchez-Vives & Slater, 2005; Mestre, 2015). Notably, Tromp and colleagues first coupled electroencephalography (EEG) recordings and a VR environment to examine linguistic processing (Tromp, Peeters, Meyer & Hagoort, 2018). Indeed, everyday language processing generally happens in rich, contextualized environments and, importantly, numerous extraneous factors including the interlocutors, social context and physical cues have a strong influence on how language is understood. By combining EEG and VR, Tromp and colleagues sought to overcome the limitations of impoverished linguistic contexts, by immersing participants in lifelike environments (Tromp et al., 2018).

There are few instances of such coupling, however, and previous studies of the interaction between linguistic and motor processes have used electroencephalography (EEG) in non-interactive protocols (Alemanno et al., 2012; Aravena et al., 2010; Moreno et al., 2015). In a recent study (Zappa et al., 2019), we recorded EEG signals in a situation where participants had to manipulate virtual objects, following the auditory presentation of action verbs, while immersed in a VR (CAVE) environment. Results provided clear evidence, in this naturalistic setting, of how motor and linguistic processes interact.

This setup involved two major challenges that we addressed and that we will present below. First, the continuous EEG activity needed to be accurately synchronized to the auditory and visual events presented to the participants in the virtual environments. Second, manipulation interfaces needed to be specially designed so that the participants could physically interact with virtual objects.

The present paper describes in detail the experimental setup used (anonymized), combining EEG and fully immersive VR to study the effect of premotor activity on the processing of action words via a Go-Nogo paradigm. The experimental design was kept simple to allow us to concentrate on resolving the technical complexities involved in integrating EEG in the VR environment. In the remainder of this article we will present the two principal elements of the experimental setup: the stimulation system and the acquisition system. By stimulation system, we refer to the virtual reality system, which creates the experimental environment and controls the presentation of the experimental stimuli. The acquisition system refers to the EEG system, which records the brain activity throughout the experimental session.

Participants were seated in a 3D CAVE (Cruz-Neira, Sandin & DeFanti, 1993) and heard a series of French action verbs, followed by the presentation of virtual objects (Figure 5.1). On Go trials, participants carried out the corresponding manual action on one of virtual objects, the target, presented before them. EEG was recorded continuously throughout the experimental session.



* Not included in analysis to avoid artefacts evoked by visual cue.

Figure 5.1: *Detail of a single Go and Nogo trial over time. The only difference between trial-types is that, on Go trials, participants carry out the relevant action, while on Nogo trials, they keep their hands in the neutral position.*

5.2 Methods

5.2.1 Stimulation system

Cave automatic virtual environment (CAVE)

Our CAVE is a projection-based stereoscopic immersive virtual reality system consisting of 4 projection screens in a cubic configuration; the 4 screens correspond to the front, left and right walls and the floor (Figure 5.2). In our configuration, the vertical walls are 4m high and 3m wide, the floor measures 3m². Each screen is fed by a stereoscopic image of 1600x1200 pixels at a frequency of 60Hz, projected by an active video projector. Liquid Crystal Shutter glasses (Volfoni ®) are used to deliver stereoscopic images to the participant. A motion capture system (Advanced Realtime Tracking, ArtTrack ®) captures the position and orientation of passive markers. at a frequency of 60Hz. The ArtTrack cameras are connected to a specific controller that delivers tracking data to a graphics cluster. Markers are attached to the stereo glasses, which permits the graphics system to compute and project the corresponding stereoscopic images in real-time, according to the user's viewpoint. The images are produced by the 3D engine Unity3D ®. A homemade unity asset carries out the distribution and synchronization of data between the graphics computers, ensuring the display of a coherent virtual environment on the screens. We used a specific configuration of the cameras (Figure 5.2) to optimize the capture of hand movements (see below).

Figure 5.2: *Cave automatic virtual environment (CAVE). The virtual environment represents an office, with a desk on which the participant can manipulate objects.*

Finger-tracking system

During the experiment, participants wore a finger-tracking system (add-on to the ART tracking system) on their right hand (Figure 5.3). The finger-tracking device allows one to track the orientation of the hand and the position of three fingers. The finger-tracking system is wireless and composed of four active markers: one located at the top of the hand, which captures the position and orientation of the back of the hand, and one on the tip of the thumb, the index finger and the middle finger. The finger markers are time-sequentially addressed and controlled by the active hand target. The ART controller provides the position and orientation of the hand, as well as the position, the orientation and the size of each phalange of the three fingers (thumb, forefinger, middle finger). The ring and the little fingers reproduce the same movement as the middle (tracked) finger. This system enabled the participant to manipulate the virtual objects while tracking their hand movements.

Figure 5.3: *ART finger-tracking system (from https://ar-tracking.com/)*

A specific module was developed in Unity3D to model a physical hand interacting with virtual objects. This physical hand is composed of several 3D primitive shapes: the phalanges are represented by cylinders, the joints by spheres and the edge of the palm is represented by a number of cylinders. All of these physical objects can enter into collision with a virtual object. This is computed by the physics engine of Unity 3D. Unity provides callback functions that manage contacts between physical objects. A virtual object is considered as seized when the algorithm detects that the tips of both the thumb and index or middle finger touch it simultaneously. Inversely, if the tips of the thumb and the index or middle finger are no longer in contact with the virtual object, it is released and falls. This algorithm can reproduce pushing, grasping and releasing actions on a virtual object, according to physical parameters such as the speed of the hand, its trajectory, the position of the object relative to the hand or the weight of the object.

A “neutral” hand position, in which participants placed both hands flat on the transparent table in front of them, was defined; a colored sticker marked the position of their right hand on which they wore the finger-tracker. Participants were instructed to keep their hands in this position when not performing an action. The finger-tracking system detected when, after having performed an action, the participant’s hand had returned to this position and the following experimental trial could only begin once this neutral position was detected.

During the training phase, participants learned how to perform all actions correctly. We observed how quickly the participants were able to adapt to the finger-tracking glove while executing the actions, the ease with which they carried out the actions using it and how responsive the glove was to the different movement characteristics. The number of trials needed depended on the complexity of the action. For the simplest actions such as “cacher”

[hide], no learning was necessary. In contrast, for more difficult actions such as “relever” [place the object upright] participants often repeated the action over a few trials.

The finger tracker also provided information concerning the participants’ hand movement trajectories. Such data could be used to verify, offline, the ease with which individual participants correctly executed each action on the virtual objects. In addition, one line of further investigation is the potential use of this movement data as a regressor in the analysis of the EEG signal to determine a link between certain characteristics of the action to be performed and pre-motor activity.

Plexiglas table

Participants were seated at a specially-designed Plexiglas table onto which the virtual objects were projected (Figure 5.4). Plexiglas was used as this transparent material allowed the stereo objects to appear as if they had been placed on the table (while preserving correct projection of images on the screens). In addition, this physical (tangible) object provided a haptic reference that facilitated interactions with the virtual objects, such as pushing an object across a surface, causing an object to fall, etc. Finally, this table provided a solid support upon which the participants could rest their hands when not manipulating the virtual objects, while impinging as little as possible on the virtual environment

Figure 5.4: *A participant seated at the plexiglas table manipulating a virtual object inside the CAVE system.*

5.2.2 Acquisition system: Electroencephalograpy

Electroencephalography (EEG) activity was recorded continuously throughout the experimental session by means of a 64-channel active EEG system (BioSemi Active Two) in which electrodes are arranged in the International 10-20 configuration, which specifies the location of the scalp electrodes. The Biosemi Active system is standard for non-clinical EEG acquisitions for a number of reasons, most notably the following: the electrodes used in this Biosemi system are “active”, which means that the signal is amplified at the level of the electrodes and has the advantage of keeping the output impedance of electrodes below 1W and maximizing the input impedance. The active electrode system reduces the problem of electrical interference with extraneous outputs (flat screen, projectors, lighting, etc) arising from the CAVE system. Another specificity of the Biosemi system is that the “ground” electrode is replaced by two electrodes, the Common Mode Sense (CMS) and the Driven Right Leg (DRL, a passive electrode), which form a feedback that drives the mean potential of the participant as close as possible to the reference voltage of the AD-box of the Biosemi system. Individual electrodes were adjusted to a stable offset with an upper limit of 20mV; in Biosemi the offset is measured in relation to the Common Mode Sense (CMS) electrode. Blinks and vertical eye movements (VEOG) were monitored via two external electrodes placed under each eye and horizontal eye movements (HEOG) were monitored via two electrodes positioned at the outer canthus of each eye. External electrodes were also positioned at both the left and right mastoids, which were used as references offline. EEG was recorded at a rate of 2048 Hz.

5.2.3 Synchronization of stimulation and acquisition systems

EEG allows one to capture changes in brain activity with high temporal resolution. In this study we were interested in studying the EEG activity that both preceded and followed the onset of a critical stimulus. In such event-related potential (ERP) studies we are interested in brain activity in relation to the onset of a stimulus. Accuracy at the millisecond level is necessary to reconstruct this event-related activity. Therefore, it is crucial that the precise moment of presentation of a critical stimulus in the continuously recorded EEG can be identified. In the context of the current study, this implied ensuring that the stimulation system (VR system) that controlled the time of presentation of the stimuli and the acquisition system (EEG system) that recorded the neural activity be tightly synchronized over time.

The integration of the VR and EEG systems is presented, graphically, in Figure 5.5, below. The Unity3D software controlled the timing and presentation of both auditory (the verbs) and visual stimuli: the trial-type cue (Go or Nogo), as well as the virtual objects. The auditory stimuli were sent to the loudspeakers and to the EEG system, via the

“ERGO1” (trigger) input of the Biosemi AD box, simultaneously. Thus, the acoustic stimuli were recorded onto an auxiliary EEG channel and precise temporal synchronization with the activity recorded by the electrodes was ensured.

To ensure that the precise time of presentation of the visual stimuli could be identified in the continuous EEG, we used photodiodes to mark the onset of presentation of the visual stimuli. A small, white square was projected onto the bottom left-hand corner of one of the lateral walls of the CAVE and a photodiode was positioned in the center of this square. As for the acoustic signal, the photodiode was connected to the AD box via the second ERGO2 input such that the photodiode signal could be recorded as an auxiliary EEG channel. The presentation of the visual stimuli and the change in color of the white square from white to black was triggered simultaneously. This change in luminosity was detected by the photodiode and was reflected in the photodiode signal as a step function, the onset of which also marked the moment of presentation of the visual stimulus. The luminosity change can be considered as a trigger and induced a response by the photodiode that precisely marked the onset of the visual stimuli. The photodiode also allowed us to distinguish between the different types of visual stimuli, the Go and Nogo cue, and the two virtual objects presented on the left or right. This was achieved by varying the duration of the black square: 750ms for the Go cue, 500ms for the Nogo cue, 1500ms for left-side target on left and 1000ms for right-side target. This allowed us to determine offline, from the photodiode’s behavior, not only the time of presentation of the visual stimulus both also the stimulus type.

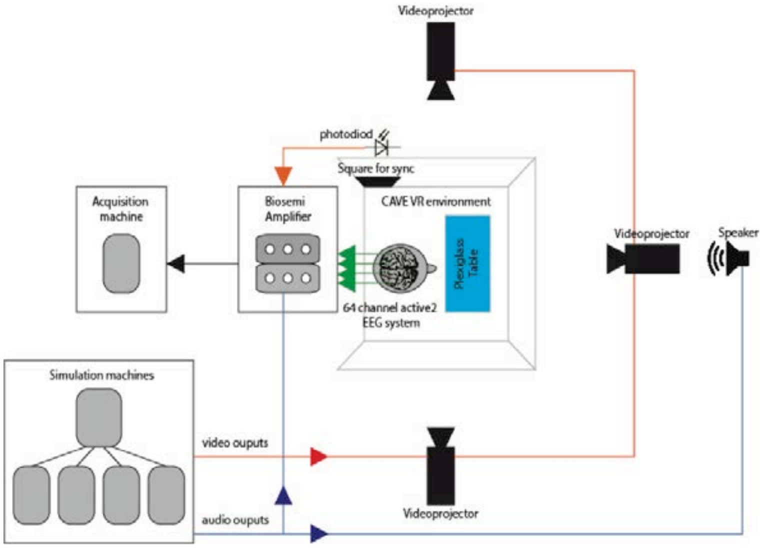


Figure 5.5: *Graphic summary of the integration of the stimulation (VR) and acquisition (EEG) systems.*

Auditory Stimuli

To extract the precise onset time of presentation of each acoustic stimulus, the action words, we carried out a cross-correlation between the continuous auditory signal recorded onto the auxiliary EEG channel and the signal of each individual action word in turn and then calculated the lag of the maximum correlation value. This yielded both the onset time of every instance of each individual action verb and identified that action word in the continuous auditory signal. However, it was necessary to take into account the difference in sampling rate between the original auditory stimuli, sampled at 48kHz and the auditory signal recorded onto the auxiliary EEG, which was sampled at 2048Hz, the default Biosemi sampling rate. The individual auditory stimuli were, therefore, down sampled to 2048Hz before carrying out the cross-correlation.

Visual Stimuli

To identify the type of visual stimulus from the photodiode signal, it was necessary to extract the onset and offset times of each negative deflection. To this end, the photodiode signal was first detrended, half-wave rectified and lowpass filtered at 8Hz, to remove an observed oscillation at this frequency. The response of the photodiode to each decrease in luminosity was converted to a clean step function by setting all activity lower than the mean activity to zero and all activity greater than zero to one. Therefore, a simple first order difference was applied to determine the onset (+1) and offset (-1) times.

EEG quality check

Prior to the experiment, it was important to assess the quality of the EEG signal in the CAVE and compare it to normal experimental recording conditions, in a Faraday cage. This allowed us to evaluate the extra noise present in our data due to a less controlled environment, which was particularly important as we intended to analyze the frequency content of the EEG signals. Figure 5.6 presents the power spectral density (PSD) of resting-state EEG in a usual experimental setup (Faraday cage), using the multi-taper method (Thomson, 1982), and in our VR setup, with all systems turned on (projectors, motion tracking cameras, finger tracking, VR glasses, and the participant executing a simple arm movement). EEG activity was captured from the Fz frontal electrode, for 2 seconds.

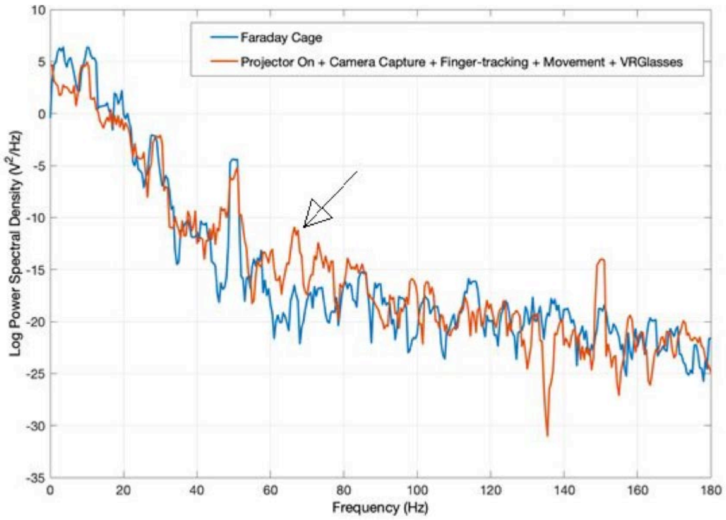


Figure 5.6: *Power spectral density of 2seconds of EEG activity captured by the frontal electrode in laboratory conditions (Faraday cage) and in our VR system, with all devices turned on. The arrow points to the VR trace.*

Offline detection of even onsets

A comparison of the spectra (Figure 5.6) of both recording conditions (CAVE and Faraday cage) revealed comparable line noise (50Hz). Indeed, in general, the signal recorded in the VR environment presents slightly greater energy in the higher frequency range (above 50Hz). In particular, we can observe a peak around 60Hz (see arrow in Figure 5.6), which is likely due to the synchronization between the projectors and the stereo glasses.

5.3 Conclusion

We presented here a novel method to synchronize EEG recording, sensorial (visual and auditory) rendering and motor actions recording in an immersive virtual environment (CAVE setup). Our protocol allowed us to successfully synchronize the EEG and CAVE stimulation and acquisition systems, whereby the EEG data was perfectly time-locked to stimulus presentation in the CAVE. The detailed results of the experiment using this method are reported in (anonymous). This is a demonstration of how EEG can be successfully combined with an interactive virtual environment, which should open avenues for future studies, analyzing further the interactions between motor behavior, linguistic processing and immersion.

References

Alemanno, F., Houdayer, E., Cursi, M., Velikova, S., Tettamanti, M., Comi, G., Cappa, S. F., & Leocani, L. (2012). Action-related semantic content and negation polarity modulate motor areas during sentence reading: An event-related desynchronization study. *Brain Research*, *1484*, 39–49.

Aravena, P., Hurtado, E., Riveros, R., Cardona, J. F., Manes, F., & Ibáñez, A. (2010). Applauding with closed hands: Neural signature of action-sentence compatibility effects. *PloS One*, *5*(7), e11751.

Cruz-Neira, C., Sandin, D. & DeFanti, T. (1993). Surround-screen projection-based virtual reality: The design and implementation of the CAVE, *Proceedings of SIGGRAPH' 93*, 135-142.

Loomis, J.M., Blascovich, J.J. & Beall, A.C. (1999). Immersive virtual environment technology as a basic research tool in psychology. *Behavior Research Methods*, *Instruments and Computers, 31,* 557-564.

Mestre, D. R. (2015). On the usefulness of the concept of presence in virtual reality applications. *The Engineering Reality of Virtual Reality 2015*, *9392*, 93920J.

Moreno, I., de Vega, M., León, I., Bastiaansen, M., Glen Lewis, A., & Magyari, L. (2015). Brain dynamics in the comprehension of action-related language. A time-frequency analysis of mu rhythms. *NeuroImage*, *109*, 50–62.

Peeters, D. (2019). Virtual reality: A game-changing method for the language sciences. *Psychonomic Bulletin & Review*, *26*(3), 894–900.

Sanchez-Vives, M. & Slater, M. (2005). From presence to consciousness through virtual reality. *Nature reviews. Neuroscience, 6*. 332-339.

Tarr, M.J. & Warren, W.H. (2002). Virtual reality in behavioral neuroscience and beyond. *Nature Neuroscience, 5*, 1089-1092.

Thomson, D.J. (1982). Spectrum estimation and harmonic analysis. *Proceedings of the IEEE*, vol. 70, no. 9, pp. 1055-1096.

Tromp, J., Peeters, D., Meyer, A. S., & Hagoort, P. (2018). The combined use of virtual reality and EEG to study language processing in naturalistic environments. *Behavior Research Methods*, *50*(2), 862–869.

Chapter 6

The neural correlates of embodied L2 learning

Zappa, A., Bolger, D., Pergandi, J.-M., Fargier, R., Mestre, D., & Frenck-Mestre, C., submitted to *Neurobiology of Language*

The present study aims to investigate how naturalistic actions in a highly immersive, multimodal, interactive 3D virtual reality (VR) environment may enhance word encoding by electroencephalography (EEG) recording in a pre/ post-test learning paradigm. Both imaging and electrophysiological data have established motor activation during language processing, and behavioral data has shown that coupling word encoding with gestures enhances learning. However, the neural underpinnings of facilitated action language learning have yet to be elucidated. Herein, we couple EEG recording with virtual reality to examine whether "embodied learning", or learning that occurs using specific physical movements that are coherent with the meaning of new verbs, creates linguistic representations that produce greater motor resonance (mu and beta event-related synchronization/desynchronization), due to stronger motor traces, compared to learning without accompanying specific gestures. We will also investigate whether greater motor resonance while listening to learned action verbs post-learning correlates with improved retention.

6.1 Introduction

Following decades of considering language processing as independent from motor processes, theories of embodied cognition are hard at work building an argument for reuniting the two (Barsalou, 2008). Generally speaking, these theories propose that cognition is grounded in multimodal representations originating in human experience (Pulvermüller, 2005). As concerns language, amodal symbolic linguistic representations (Fodor, 1983) are replaced with modal representations, giving motor processes an essential role in language processing (Wilson & Golonka, 2013). Indeed, sensory and motor systems are recruited during lexical processing, both during development (James & Swain, 2011) and in adults (Hauk, Johnsrude & Pulvermüller, 2004; Pulvermüller, 1999, 2005; Thompson-Schill, 2003). Furthermore, neuroimaging has revealed an overlap in neural mechanisms for processing speech and hand movement (Nishitani, Schürmann, Amunts & Hari, 2005). Along the same lines, gestural studies have suggested that gesture and speech comprise an integrated system (Goldin-Meadow, 2011; Graziano & Gullberg, 2018). Language processing is facilitated by gesture such that language-action congruency that occurs early in sentence processing can facilitate lexical retrieval (Glenberg & Kaschak, 2002). Conversely, the retrieval of stored semantic representations directly influences sensorimotor activation as indexed by greater motor preparation when congruent action language is presented prior to movement (Aravena, 2012). Importantly, incongruity between actions and meaning can cause interference in meaning retrieval (Aravena et al., 2010; Barsalou, 1999). These results suggest that motor representations are not simply reactivated by linguistic representations post-lexically (Mahon & Caramazza, 2009) but that they can play a role in representations. In the current investigation, we will examine the interaction between motor and semantic processes and how it affects the mapping of novel action verbs to physical actions. Word encoding will be coupled with compatible physical actions in an interactive virtual environment that enables pseudo-natural movements, to test whether motor activation enhances novel action-word learning in a foreign language. We will examine whether action verbs learned with specific actions produce greater motor activation post-training, as revealed by beta and mu event-related synchronization/desynchronization, compared to verbs learned without accompanying actions. Finally, we will examine whether our experimental manipulation leads to improved retention, due to a stronger motor trace in memory (Engelkamp & Krumnacker, 1980).

Encoding new words is an essential part of language acquisition and second-language (L2) learning. Word learning studies on infants and children have reported important findings on questions such as the rapidity of object-word mappings (Friedrich & Friederici, 2008) and the learning outcome of massed versus distributed presentation (Childers & Tomasello, 2002). Given that the majority of the words we know are learned after childhood (Borovsky, Elman, Kutas & 2012), and that important differences exist in how words are encoded during childhood versus adulthood (Borovsky, Elman & Kutas, 2012), adult word encoding deserves special attention. Novel word or L2 word learning studies have investigated how adults learn new words. In a seminal study McLaughlin and colleagues (2004) found differences in learners' cortical activity when processing pseudowords compared to L2 words to which they had been exposed, as indexed by an N400 effect, after around 14h of classroom instruction (McLaughlin, Osterhout & Kim, 2004). After around 63 hours of instruction, they also found semantic effects for L2 words versus pseudowords. Similarly, differences in electrophysiological indices of word processing between the L1 and the L2 were seen via an N400 increase for L2 words after one semester of learning (Soskey, Holcomb & Midgley, 2016). The former studies, however, investigated extended L2 training and did not allow to for the observation of cortical changes occurring during the very first moments of encoding.

The benefit of physical movement for language learning and memorization is not a new topic (Moskowitz, 1976; Allen, 1995). Behavioral studies dating back to the 1980s have shown that illustrative gestures support language retention better than other conditions (Engelkamp & Krumnacker, 1980; Engelkamp & Zimmer, 1984). For example, Engelkamp and Krumnacker (1980) showed that verb phrases such as “shuffle the cards” were better memorized when learners performed representative gestures during learning compared to either watching someone else perform the action, imagining the action or simply listening to the sentence. Outside of the language domain, a number of studies have shown that participants encode new information better when they perform gestures that are congruent with the new content. Physical activity facilitated the integration of sung melodies (Wakefield & James, 2011) as well as mathematical (Kontra, Lyons, Fischer & Beilock, 2015) and scientific principals (Johnson-Glenberg & Megowan-Romanowicz, 2017; Johnson-Glenberg, Megowan-Romanowicz, Birchfield & Savio-Ramos, 2016) more than verbalization. For instance, children who were asked to gesture while learning a new mathematical concept were better at retaining what they had learned over time as opposed to those who were simply asked to verbalize the new information (Cook, Mitchell & Godin-Meadow, 2008).

This “enactment effect”, has been replicated in studies focusing on second-language or artificial language word-learning studies. After 20-30 minutes of leaning novel words by simply pointing or touching the corresponding objects, participants showed associations between sensorimotor experiences (the location of an object in a vertical space) from training and the novel words (Öttl, Dudschig & Kaup, 2016). Importantly, it has been argued that truly embodied learning involves “self-performed” or “self-generated” action, as opposed to simply observing or imagining action (James & Bose, 2011; James & Swain, 2011; Johnson-Glenberg, 2017, 2018; Johnson-Glenberg & Megowan-Romanowicz, 2017). Two recent taxonomies for embodied learning propose different levels of embodiment for education based on how physically engaged learners are, as well as the congruency between gestures and the content being learned (Johnson-Glenberg & Megowan-Romanowicz, 2017; Skulmowski & Rey, 2018). In other words, highly embodied learning generally implies that learners physically perform gestures or movements that are directly linked to the content they are learning (Johnson-Glenberg, 2018). Both L1 and L2 lexical encoding studies generally use representative or iconic gestures (McNeil, 1992), which illustrate and map onto meaning directly. For instance, children who performed representative gestures while learning L2 lexical items were better at producing them when tested than those who learned with images (Tellier, 2008). Similarly, adults were better at recalling L2 words encoded audio-visually with accompanying representative gestures as opposed to those encoded only audio-visually (Macedonia & Knösche, 2011). In like fashion, de Nooijer and colleagues found that words describing object manipulation (but not abstract words or words describing locomotion) were better learned when participants were asked to imitate a gesture during encoding and retrieval, than when participants did not gesture (de Nooijer, van Gog, Paas & Zwaan, 2013). This result held true in a delayed post-test, one week after training.

The studies cited above indicate that action boosts memory performance and therefore supports language encoding. However, it is still unclear what cognitive processes are behind this facilitation. One explanation is that physical action relays and helps establish implicit knowledge. Indeed, we often express information of which we are not consciously aware through gestures (Church & Goldin-Meadow, 1986). Broaders and colleagues (2007) found that children who were encouraged to gesture while solving math problems were more receptive to instruction during subsequent math lessons and showed enhanced learning compared to children told not to gesture (Broaders, Cook, Mitchell & Goldin-Meadow, 2007). According to Sun and colleagues (2001) combining explicit and implicit performance can aid in learning new skills (Sun, Merrill & Peterson, 2001). The multisensory learning theory (Shams & Seitz, 2008) promotes a different approach to the gesture-learning facilitation. It suggests that natural everyday learning involves multiple senses, which leads to better learning outcomes than single-sense learning (for instance auditory). Evidence for this was provided in an L2 learning paradigm that manipulated presentation format and gesture. Learning outcomes two and then six months post-training were best for words learned with congruent gestures. Purely auditory learning yielded the lowest performance (Mayer, Yildiz, Macedonia & von Kriegstein, 2015).

According to the theory of Hebbian associative learning, the synchronous activity of neurons forms neuronal assemblies (Hebb, 1949); hence when lexical items are acquired along with action, cortical areas involved in language processing and those involved in action planning and execution quickly develop into shared neural circuits (Pulvermüller, 1999; 2005; Tomasello, Garagnani, Wennekers & Pulvermüller, 2018). To better understand how learning may be enhanced by movement, several studies have examined the neural underpinnings of lexical-motor interactions. Results from a functional magnetic resonance imaging (fMRI) study conducted with children revealed that when listening to learned words, the motor system was only recruited for items learned through direct interaction compared to simple observation (James & Swain, 2011). In an fMRI study, Macedonia, Müller and Friederici (2011) found that novel words learned by watching and performing iconic gestures, as opposed to meaningless gestures, were better retained and produced activity in the premotor context. Finally, Krönke, Mueller, Friederici and Obrig (2013) found no behavioral differences in novel word leaning whether training involved iconic gestures, grooming gestures or no gestures at all. However, fMRI results revealed larger activation in the semantic network or "deeper semantic encoding" for words learned with iconic gestures.

Despite the importance of the above studies, fMRI may not be the ideal tool to show motor to language effects or vice versa. Indeed, much debate surrounds the role of motor activation during language processing. One of the arguments against embodied semantics is that language-induced motor activations are post-lexical and not a necessary part of language processing (Mahon & Caramazza, 2008). High temporal resolution — an advantage of electroencephalography (EEG) compared to fMRI — is hence an important element in arguing for embodied language representations. EEG can be used to quantify motor cortex activity by measuring event-related synchronization/desynchronization via stimulus-locked time-frequency analysis (Vukovk & Shtyrov, 2014). Specifically, event-related desynchronization (ERD), or suppression, in the mu (8-13 Hz) and beta (13-30 Hz) bands has been associated with sensorimotor activation involved in movement preparation and execution (Pfurtscheller & Lopes da Silva, 1999; Niccolai et al., 2014; Pineda, 2005). A decrease in the alpha rhythm has likewise been linked to motor imagery (Holler et al., 2013).

Recently, desynchronization in oscillations associated with motor processes has also been observed during action language understanding. Reading sentences describing manual actions versus abstract sentences led to a suppression of mu rhythms at fronto-central sites (Alemanno et al., 2012; Moreno, de Vega, León, Bastiaanse & Magyari, 2015). To our knowledge, the only study that has used time-frequency to measure motor activation during language processing pre and post-training was conducted by Fargier and colleagues (2012). They showed that learning novel words in association with specific self-performed actions led to greater mu desynchronization post-training, hence motor activation, compared to learning in association with abstract animations. (Fargier et al., 2012). However, on the second day of training, a fronto-central distribution of the effect, as opposed to a typical central parietal mu distribution, lead the authors to conclude that it was confined to a convergence zone. Embodied cognition binds social and physical contexts to cognition, and therefore the environment in which learning takes place could potentially play an important role in learning outcome (Black, Segal, Vitale & Fadjo, 2012).

According to Atkinson (2010), learning is not just a mental process but one that occurs in environments made up of “bodies, cognitive tools, social practices and environmental features” and this multimodality calls for an experimental approach that is likewise multimodal. One caveat of experimental protocols that examine “embodied” L2 learning is that they often impose strict physical and environmental limitations. Given the need for control, movement is generally reduced to minimal hand actions and training most often occurs in isolated and decontextualized environments (Peeters, 2019). This is especially true of studies that analyze the neural correlates of language processing and learning using brain-imaging techniques such as fMRI or magnetoencephalography (MEG). For instance, when interaction with objects has been made possible, it has been limited to pointing at or touching objects, hence making it impossible for participants to map specific actions to specific words. When one considers the importance of interlocutors, social context and physical cues on how language is understood in real-life (Knoeferle, 2015), physical and environmental limitations likely affect how language is learned. Within the framework of embodied cognition, it is especially important to take a closer look at the gap between real-life language processing and that which takes place in an experimental environment (Tromp, Peeters, Meyer & Hagoort, 2018).

Virtual reality (VR), which aims to replicate reality through three-dimensional, life-like images and sounds, represents an important tool for investigating embodied language learning. Numerous L2 studies have used VR paradigms involving varying degrees of immersion to investigate language learning. These studies have generally found facilitation for learning in immersed conditions compared to word-word or picture-word paired association (Berns, Gonzalez-Pardo & Camacho, 2013; Lan, Fang, Legault & Li, 2015). On the low-immersion end of the spectrum lie studies using gaming and social network platforms in which participants interact with object via avatars as a virtual environment learning condition (for a review, see Legault et al., 2019). Participants who learned L2 words in a *Second Life* virtual environment required only half as many repetitions to integrate the new words as those who learned through picture-word associations (Lan et al., 2014). Furthermore, they showed neural activations that were more distributed and associated with embodied networks compared to the control group. However, *Second Life* paradigms are limited when it comes to exploring truly interactive embodied learning as they rely on a desktop computer and a mouse, which restrict both visual perception and physical interaction. In order to investigate which specific elements of interactive VR environments lead to L2 integration, as well as individual differences in L2 learning, Legault and colleagues (2019) taught participants a set of L2 words using an ecologically valid immersive virtual reality zoo or kitchen, with word-word paired association as a control. Participants — especially less successful learners — showed higher accuracy in the immersive VR condition (Legault et al., 2019).

Peeters (2019) claims that VR “shifts the theoretical focus toward the interplay between different modalities [...] in dynamic and communicative environments, complementing studies that focus on one modality in isolation.” (p.1, 2019). Immersive, 3-D environments use visual and auditory stimuli to create sensory illusion, providing participants with believable environments (for a review see Bohil, Alicea & Biocca, 2011). Further, participants’ head and body movements are tracked by input tools (e.g. hand controls) and they are given real-time feedback for their actions, which provides a sensation analogous to real life (Burdea & Coiffet, 2003). The fact that participants can interact with the environment by manipulating virtual objects and carrying out naturalistic actions gives them a sense of “agency” (Johnson-Glenberg, 2018). Compared to traditional experiments, this leads to the sensorimotor system being more implicated and responses and actions being closer to what occurs in real life (Bohil et al., 2011). Finally, VR combines ecological validity with full control over the onset, location and duration of presentation of the multimodal stimuli. Very few studies have combined virtual reality and EEG to study language processing. In an exploratory EEG-VR experiment, participants listened to a sentence (“I just ordered this salmon”) and saw a virtual object that either matched (salmon) or mismatched (pasta) the object in the sentence. An N400 effect was observed for mismatched versus matched pairs, and the authors interpreted this a proof of validity for combining VR and EEG for testing language processing (Tromp et al., 2018). Recently, motor-related EEG activity was measured in an interactive virtual reality environment while participants performed a Go-Nogo task and listened to action verbs prior to executing the corresponding actions. Motor activation was found through mu and beta event-related desynchronization (ERD) for both Go and Nogo trials, during verb processing and prior to movement proper, providing compelling evidence in a naturalistic setting of how motor and linguistic processes interact (Zappa et al., 2019). Moreover, greater ERD was found for Go trials, suggesting that motor preparation influenced semantic processing. These results provide the basis for the present study, investigating the mapping of new linguistic labels to motor actions.

Our study will use a combined EEG-VR methodology to explore the neural correlates of embodied L2 learning. Unlike the two EEG-VR studies described above, EEG will not be recorded directly within the VR environment but pre and post VR training. Using a head mounted VR system (Oculus Rift) and controller, participants will learn an auditory L2 lexicon of action verbs by mapping words to congruent self-performed physical actions. The type of motor action will correspond to the specific action (picking up an object throwing it for the verb “throw”). A control group will learn the same L2 lexicon but will instead perform a single action for every verb: pointing to the object. Both pre and post-training, learners’ knowledge of the semantic meaning of the training verbs will be measured behaviorally and through EEG using a match-mismatch task. Motor resonance will also be measured using

EEG while participants listen to the training verbs as well as filler verbs, both pre and post-training. We expect motor resonance during auditory verb processing post-training to vary as a function of learning condition. We hypothesize that representations of verbs learned with specific actions will carry a stronger motor trace and hence produce greater motor activation than verbs learned in the pointing condition. We also predict that embodied learning using specific self-performed congruent physical actions will lead to better learning outcomes post-training compared to the pointing control condition.

6.1.1 Hypotheses

1

In accordance with the theory that learning lexical items along with action can form shared neural networks (Pulvermüller, 1999; 2005; Tomasello et al., 2018) and studies showing greater motor activation for object labels learned with direct object interactions (James & Swain, 2011) or specific self-performed actions (Fargier et al., 2012), we expect to find greater beta (13-30 Hz) and mu (8-12 Hz) desynchronization (motor activation) post-training compared to pre-training during the processing of the training verbs (passive listening task). Given that only training verbs will have been associated to meaning, these effects are expected to be greater for training verbs compared to filler verbs.

2

Activity in the premotor context has been found when learners process verbs learned with iconic gestures but not those learned with meaningless gestures (Macedonia et al., 2011). We therefore expect to find greater motor resonance for verbs learned in the Specific action condition compared to the Pointing condition.

3

An N400 effect for mismatched compared to matched visual image-auditory word pairs has been associated with difficulty in semantic integration (Kutas & Federmeier, 2011). During the match-mismatch task, we expect that, pre-training, we will not find an N400 effect for match versus mismatch trials. Post-training, we expect to find greater N400 amplitude for mismatch versus match trials in both conditions, due to participants accessing the semantic meaning of newly learned verbs.

4

Along with studies in non-linguistic domains showing enhanced learning when gestures are used (Broaders et al, 2007; Cook et al., 2008; Johnson-Glenberg et al., 2016; Johnson-Glenberg & Megowan-Romanowicz, 2017; Kontra et al., 2015; Wakefield & James, 2011), both behavioral (Mayer et al., 2015; Tellier, 2008) and electrophysiological (Macedonia & Knösche, 2011; Macedonia et al., 2011; de Nooijer et al., 2013) evidence from L2 learning studies has revealed that congruent gestures support linguistic memory and encoding and improve performance in the L2. We therefore hypothesize that the N400 effect outlined in hypothesis 3 will be greater for the Specific action condition compared to the Pointing condition.

5

In accordance with hypotheses 3 and 4, we expect to find a correlation between greater motor resonance during the passive listening task and a greater N400 amplitude for mismatch versus match trials in the match-mismatch task.

6

In accordance with hypothesis 4, we predict that our behavioral results will show greater accuracy for verbs learned in the Specific action condition compared to the Pointing condition.

6.2 Methods

In the current study we will manipulate the type of action performed (specific object manipulation vs pointing) during L2 learning in a VR environment. During learning, participants will visualize movements performed by a virtual hand. The Specific action group will reproduce the movement on a virtual object and the control group will point to the virtual object on which the action was performed. EEG will be recorded both pre- and post-training.

6.2.1 Ethics

This research complies with all relevant ethical regulations and has been approved by the local university ethics committee.

6.2.2 Statistical Power Analysis

For hypothesis 1, a statistical power analysis was performed for sample size estimation using G*Power 3.1 (Faul, Erdfelder, Buchner & Lang, 2009). The analysis was based on data from the published study Motor resonance during linguistic processing as shown by EEG in a naturalistic VR environment (Zappa et al., 2019) (N=20), comparing mu and beta desynchronization for Nogo vs Go trials. The effect size (ES) in this study was .8, considered to be large using Cohen's (1988) criteria. However, for the power analysis we used a "medium" effect size, as effect sizes based on previous results are often overestimated (Gelman & Carlin, 2014). With an alpha = .05 and power = 0.80, the projected sample size needed with this effect size is approximately N = 34 for this simplest within group comparison. Thus, our proposed sample size of 40 will be adequate for the main time-frequency related hypothesis of this study and should also allow for expected attrition. We also expect that this sample size will be adequate for our five other hypotheses as they either concern mu and beta desynchronization as well or an N400 effect, which has been shown to require a smaller sample than 40 (Fields & Kuperberg, 2019).

6.2.3 Participants

Eighty (40 per group) right-handed French native speakers (40 women, aged 20–26) with no previous knowledge of Serbian or related languages will participate in the study. Participants will be right-handed volunteers from the student population of the Aix-Marseille Université, having stereoscopic vision and no history of neurological insult. All participants will give their written informed consent prior to the experiment. Participants will receive 40 euros for their participation.

6.2.4 Stimuli

Auditory stimuli consist of 12 imperative transitive verbs in Serbian that are not transparent with their translation equivalents in French, Spanish, Italian, Portuguese, German or English. Serbian is a South Slavic language that is linguistically distant from both Latin and Germanic languages such that transparency poses little threat. Auditory verbs were recorded in a professional sound booth and produced by two trained female native speakers. The verbs denote actions that can be performed using one's hand and arm, and were previously validated in a VR environment (Zappa et al., 2019): /gurni/[push], /zagrɛbi/ [scratch], /pusti/ [drop], /batsi/ [throw], /okrɛni/ [pivot], /prɛmɛsti/ [move], /kutsni/ [tap], /uvati/ [catch], /podini/ [lift], /lupi/ [hit], /obori/ [tip over], /protrɛsi/ [shake].

One set of verbs, recorded by speaker 1, will be used for learning. A second set of auditory verbs, recorded by speaker 2, will be used for the two EEG tasks. A set of 12 filler verbs denoting different actions was recorded by speaker 2, for the passive listening EEG task. Visual stimuli for learning will consist of an office environment containing a 3D 10-point star polygon and a CRT screen (Figures 6.1 & 6.2)

Animations of hand and arm movements corresponding to the verbs above, performed on the 3D 10-point star polygon, were recorded. These animations will be used in both learning conditions to teach participants the movements that correspond to the verbs. They will also be used for the match-mismatch task pre and post-training.

Figure 6.1: *VR environment and CRT screen*

Figure 6.2: *CRT screen*

6.2.5 Learning apparatus

An Oculus VR headset and controller will be used for training purposes. The Oculus headset visually immerses participants by presenting them with a 360-degree visual scene and 3D virtual objects. The controller allows participants to manipulate objects while motion capture is recorded online. The headphones integrated into the Oculus headset will be used for auditory verb presentation.

6.2.6 Software

During pre and post-tests, StimPres (Tufts University) will be used for stimulus presentation on a desktop computer and a 64-channel Biosemi system (Actiview) will be used for acquisition. UNITY software will control virtual object presentation during learning.

6.2.7 General Procedure

The experiment will take place over two days. On the first day participants will undergo an EEG pre-test followed by a VR learning session. On the second day, they will participate in a second VR learning session with the same materials as day 1, followed by an EEG post-test.

6.2.8 Learning Procedure

Participants will be comfortably seated at a desk wearing a VR Oculus headset and holding a controller. Participants in both the Specific action and the Pointing conditions will be presented with an auditory verb, followed by an overt oral repetition prior to observing an action on the virtual CRT screen within the VR environment. Following this, a virtual object will appear on the virtual desk. The Specific action group will manipulate the object, performing the action observed on the virtual CRT screen (Figures 6.1 & 6.2). The Pointing group will point to the object.

Figure 6.3: *Above: a participant wears an Oculus headset and performs the verb "gurni" [push]. Below: the participant's movements translate into the virtual hand pushing the virtual object away.*

6.2.9 EEG procedure

EEG will be recorded during both pre and post-tests. Participants will be comfortably seated at a desk situated 60 cm away from a computer screen in an electrically shielded sound attenuated booth.

6.2.10 Passive listening task

During the first task participants will be asked to listen to the list of verbs passively, with no associated task. They will hear the 12 verbs used for learning and 12 filler verbs, twice. A trial will begin with an ocular fixation cross displayed in the center of the computer monitor for 200 msec prior to and for the duration of the auditory word, which will be presented via electrically shielded speakers. A visual "blink" prompt will be displayed immediately thereafter for 2 seconds. The experimental session will last roughly 15 minutes.

6.2.11 Match-Mismatch task

During the match/mismatch task the auditory verbs used in learning will be preceded by either the compatible (match) or an incompatible (mismatch) animation. A question mark will appear directly following the auditory verb. Participants will be required to answer yes or no on a response box. A visual "blink" prompt will be displayed immediately thereafter for 2 seconds. The experimental session will last roughly 25 minutes, including one break.

6.2.12 Behavioral procedure

Word retention will be tested behaviorally using a written translation task during which participants will listen to the new verbs and be asked to write down the equivalent in French, both pre and post-learning.

6.2.13 EEG data acquisition

During pre and post-tests electroencephalographic (EEG) activity will be recorded continuously from 64 scalp electrodes located at left and right hemisphere positions over frontal, central, parietal, occipital, and temporal areas by means of a 64-channel electrode cap mounted with silver-chloride active electrodes (BioSemi Active Two system AD box). Individual electrodes will be adjusted to a stable offset lower than 20µV and data will be sampled online at 512 Hz. Blinks and vertical eye movements will be monitored via an electrode placed under the right eye and horizontal eye movements will be monitored via an electrode placed at the outer canthus of the left eye. One electrode will be placed over each mastoid. EEG will be recorded continuously during the experiment and periods spanning from -100 pre-stimulus onset to 1100 msec post-stimulus onset will be used post-recording for analyses.

6.2.14 EEG data processing

EEGLAB (Delorme and Makeig, 2004) will be used to pre-process raw data. EEG data will be bandpass filtered between 0.3Hz and 80Hz. The filtered data will be re-referenced offline to the average of the two mastoids. Noisy electrodes will be determined by calculating a robust noise adjusted z-score for each, as implemented in the PREP pipeline (Bigdely-Shamlo, Mullen, Kothe, Su & Robbins, 2015). This method calculates the ratio of the median absolute deviation of high frequency components (>50Hz) to low frequency components (<50Hz) for each electrode, expressed as a z-score relative to all other electrodes. Those electrodes with a robust z-score exceeding 5 will be marked for possible rejection. This will be complemented by visual examination of the power spectral density of each electrode to determine those with excessive low and high frequency activity or contaminated by line noise.

Noisy electrodes marked for rejection will be removed. Before carrying out independent components analysis (ICA), to correct for ocular movements, sections of the EEG signal that are highly contaminated with noise will be removed from the dataset. ICA will be carried out on the continuous data of each participant using the infomax algorithm (Bell & Sejnowski, 1997) implemented in EEGLAB. Principle components analysis (PCA) will be applied prior to ICA computation to reduce the dimension of the data and speed up the ICA computation time. The number of PCA components will be estimated by calculating the explained variance of each principal component and conserving only those principal components explaining 99% of the variance. Those independent components corresponding to eye-blinks will be identified automatically using the ADJUST Toolbox (Mognon, Jovicich, Bruzzone & Buiatti, 2011) and rejected. Once the ocular artifacts are corrected using ICA, the rejected electrodes will be interpolated using spherical spline interpolation. The data will be the segmented and epochs will be visually inspected. Those contaminated by noise will be removed. The epoched data will then be divided into separate conditions for analysis. The number of trials per condition will be kept as uniform as possible across conditions. If the percentage of rejected trails exceeds 20% for any given participant, their data will be excluded.

6.2.15 EEG data analysis

ERPs

To determine the time windows in which a significant difference between match and mismatch trials emerged, a permutation test with false discovery rate (FDR) correction will be carried out on all time-points of the post-stimulus interval for each electrode; to carry out the permutation, 1000 random partitions will be performed. A significant difference will only be taken into consideration (q .05) if its duration exceeds 10 msec (5 consecutive time samples for a sampling frequency of 512Hz). The ERP (event-related potential) data will then be modeled independently in linear mixed effect models for the mean voltage amplitudes in the time windows established using permutation tests, time-locked to the onset of the verb. Analyses for the N400 component will be conducted on the data acquired at 35 electrodes, including 5 over midline (Fz, FCz Cz, CPz, Pz), and 30 lateral electrodes divided equally over the left (F1, F3, F5, FC1, FC3, FC5, C1, C3, C5, CP1, CP3, CP5, P1, P3, P5) and right (F2, F4, F6, FC2, FC4, FC6, C2, C4, C6, CP2, CP4, CP6, P2, P4, P6) hemispheres.

Event-related Spectral Perturbation (ERSP)

The event-related spectral perturbation (ERSP) will be calculated on the data from the passive listening task, time-locked to the onset of the verb, using the FieldTrip toolbox (Oostenveld, Fries, Maris & Schoffelen, 2010). To compute the ERSP, time-frequency decomposition will be effectuated at the single trial level for each participant and each condition (Pre-training, Post-training, Training verbs, Filler verbs) by applying complex Morlet wavelets over the 4Hz to 35Hz frequency band; the number of wavelet cycles will be adjusted to ensure optimal time and frequency tradeoff for the frequency bands of interest (principally mu and beta bands). The trial-level power will be averaged to yield a grand-average time-frequency map for each participant. The grand-average post-stimulus power will be expressed in terms of a change, in decibels, relative to the pre-stimulus interval (-200 msec – 0 msec prior to verb onset).

6.2.16 Timeline

We predict that the study will take seven months total to complete with the following breakdown: four months for data collection, two months for analysis and one month for the write-up.

References

Alemanno, F., Houdayer, E., Cursi, M., Velikova, S., Tettamanti, M., Comi, G., Cappa, S. F., & Leocani, L. (2012). Action-related semantic content and negation polarity modulate motor areas during sentence reading: An event-related desynchronization study. *Brain Research*, *1484*, 39–49.

Allen, L. Q. (1995). The Effects of Emblematic Gestures on the Development and Access of Mental Representations of French Expressions. *The Modern Language Journal*, *79*(4), 521–529.

Aravena, P., Delevoye-Turrell, Y., Deprez, V., Cheylus, A., Paulignan, Y., Frak, V., & Nazir, T. (2012). Grip Force Reveals the Context Sensitivity of Language-Induced Motor Activity during "Action Words" Processing: Evidence from Sentential Negation. *PLOS ONE*, *7*(12), e50287.

Aravena, P., Hurtado, E., Riveros, R., Cardona, J. F., Manes, F., & Ibáñez, A. (2010). Applauding with closed hands: Neural signature of action-sentence compatibility effects. *PloS One*, *5*(7), e11751.

Atkinson, D. (2010). Extended, Embodied Cognition and Second Language Acquisition. *Applied Linguistics*, *31*(5), 599–622.

Bakker, I., Takashima, A., van Hell, J. G., Janzen, G., & McQueen, J. M. (2015). Tracking lexical consolidation with ERPs: Lexical and semantic-priming effects on N400 and LPC responses to newly-learned words. *Neuropsychologia*, *79*, 33–41.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(4), 577–660.

Barsalou, L. W. (2008). Grounded Cognition. *Annual Review of Psychology*, *59*(1), 617–645.

Bell, A. J., & Sejnowski, T. J. (1997). The "independent components" of natural scenes are edge filters. *Vision Research*, *37*(23), 3327–3338.

Berns, A., Gonzalez-Pardo, A., & Camacho, D. (2013). Game-Like Language Learning in 3-D Virtual Environments. *Computers & Education*, *60*(1), 210–220.

Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K.-M., & Robbins, K. A. (2015). The PREP pipeline: Standardized preprocessing for large-scale EEG analysis. *Frontiers in Neuroinformatics*, *9*, 16.

Black, J. B., Segal, A., Vitale, J., & Fadjo, C. (2012). Embodied cognition and enhancing learning and motivation. In *Theoretical Foundations of Learning Environments* (D. Jonassen and S. Land, p. 2015).

Bohil, C. J., Alicea, B., & Biocca, F. A. (2011). Virtual reality in neuroscience research and therapy. *Nature Reviews Neuroscience*, *12*(12), 752–762.

Borovsky, A., Elman, J. L., & Kutas, M. (2012). Once is Enough: N400 Indexes Semantic Integration of Novel Word Meanings from a Single Exposure in Context. *Language Learning and Development: The Official Journal of the Society for Language Development*, *8*(3), 278–302.

Borovsky, A., Kutas, M., & Elman, J. (2010). Learning to use words: Event-related potentials index single-shot contextual word learning. *Cognition*, *116*(2), 289–296.

Broaders, S. C., Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2007). Making children gesture brings out implicit knowledge and leads to learning. *Journal of Experimental Psychology: General*, *136*(4), 539–550.

Burdea, G. C., & Coiffet, P. (2003). *Virtual Reality Technology*. John Wiley & Sons.

Carey, S., & Bartlett, E. J. (1978). *Acquiring a Single New Word*.

Childers, J. B., & Tomasello, M. (2002). Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Developmental Psychology*, *38*(6), 967–978.

Church, R. B., & Goldin-Meadow, S. (1986). The Mismatch Between Gesture and Speech as an Index of Transitional Knowledge. *Cognition*, *23*(1), 43–71.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. Second Edition. Hillsdale, NJ: *Lawrence Erlbaum Associates, Publishers*.

Cook, S. W., Mitch.ell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition*, *106*(2), 1047–1058.

de Nooijer, J. A., van Gog, T., Paas, F., & Zwaan, R. A. (2013). When left is not right: Handedness effects on learning object-manipulation words using pictures with left- or right-handed first-person perspectives. *Psychological Science*, *24*(12), 2515–2521.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21.

Fields, E. & Kuperberg, G. (2019). Having your cake and eating it too: Flexibility and power with mass univariate statistics for ERP data. *Psychophysiology*, *57*(2), e13468–e13468.

Engelkamp, J., & Krumnacker, H. (1980). Image- and motor-processes in the retention of verbal materials. [Image- and motor-processes in the retention of verbal materials.]. *Zeitschrift Für Experimentelle Und Angewandte Psychologie*, *27*(4), 511–533.

Engelkamp, J., & Zimmer, H. D. (1984). Motor programme information as a separable memory unit. *Psychological Research*, *46*(3), 283–299.

Fargier, R., Paulignan, Y., Boulenger, V., Monaghan, P., Reboul, A., & Nazir, T. A. (2012). Learning to associate novel words with motor actions: Language-induced motor activity following short training. *Cortex*, *48*(7), 888–899.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160.

Fields, E. C., & Kuperberg, G. R. (2020). Having your cake and eating it too: Flexibility and power with mass univariate statistics for ERP data. *Psychophysiology*, *57*(2), e13468.

Fodor, J. A. (1983). *The Modularity of Mind*. MIT Press.

Friedrich, M., & Friederici, A. D. (2008). Neurophysiological correlates of online word learning in 14-month-old infants. *Neuroreport*, *19*(18), 1757–1761.

Gelman, A., & Carlin, J. (2014). Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors. *Perspectives on Psychological Science*, *9*(6), 641–651.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, *9*(3), 558–565.

Goldin-Meadow, S. (2011). Learning through gesture. *WIREs Cognitive Science*, *2*(6), 595–607.

Graziano, M., & Gullberg, M. (2018). When Speech Stops, Gesture Stops: Evidence From Developmental and Crosslinguistic Comparisons. *Frontiers in Psychology*, *9*.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic Representation of Action Words in Human Motor and Premotor Cortex. *Neuron*, *41*(2), 301–307.

Hebb, D. O. (1949). *The organization of behavior; a neuropsychological theory* (pp. xix, 335). Wiley.

Höller, Y., Bergmann, J., Kronbichler, M., Crone, J. S., Schmid, E. V., Thomschewski, A., Butz, K., Schütze, V., Höller, P., & Trinka, E. (2013). Real movement vs. Motor imagery in healthy subjects. *International Journal of Psychophysiology*, *87*(1), 35–41.

James, K. H., & Bose, P. (2011). Self-generated actions during learning objects and sounds create sensori-motor systems in the developing brain. *Cognition, Brain, Behavior : An Interdisciplinary Journal*, *15*(4), 485–503.

James, K. H., & Swain, S. N. (2011). Only self-generated actions create sensori-motor systems in the developing brain. *Developmental Science*, *14*(4), 673–678.

Johnson-Glenberg, M. C. (2017). Embodied Education in Mixed and Mediated Realties. In D. Liu, C. Dede, R. Huang, & J. Richards (Eds.), *Virtual, Augmented, and Mixed Realities in Education* (pp. 193–217). Springer.

Johnson-Glenberg, M. C. (2018). Immersive VR and Education: Embodied Design Principles That Include Gesture and Hand Controls. *Frontiers in Robotics and AI*, *5*.

Johnson-Glenberg, M. C., & Megowan-Romanowicz, C. (2017). Embodied science and mixed reality: How gesture and motion capture affect physics education. *Cognitive Research: Principles and Implications*, *2*(1), 24.

Johnson-Glenberg, M. C., Megowan-Romanowicz, C., Birchfield, D. A., & Savio-Ramos, C. (2016). Effects of Embodied Learning and Digital Platform on the Retention of Physics Content: Centripetal Force. *Frontiers in Psychology*, *7*.

Knoeferle, P. (2015). Language comprehension in rich non-linguistic contexts: Combining eye-tracking and event-related brain potentials. In R. M. Willems (Ed.), *Cognitive Neuroscience of Natural Language Use* (pp. 77–100). Cambridge University Press.

Kontra, C., Lyons, D. J., Fischer, S. M., & Beilock, S. L. (2015). Physical Experience Enhances Science Learning: *Psychological Science*.

Krönke, K.-M., Mueller, K., Friederici, A. D., & Obrig, H. (2013). Learning by doing? The effect of gestures on implicit retrieval of newly acquired words. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *49*(9), 2553–2568.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology*, *62*, 621–647.

Lan, Y.-J., Fang, S.-Y., Legault, J., & Li, P. (2015). Second language acquisition of Mandarin Chinese vocabulary: Context of learning effects. *Educational Technology Research and Development*, *63*(5), 671–690.

Legault, J., Zhao, J., Chi, Y.-A., Chen, W., Klippel, A., & Li, P. (2019). Immersive Virtual Reality as an Effective Tool for Second Language Vocabulary Learning. *Languages*, *4*(1), 13.

Macedonia, M., & Knösche, T. R. (2011). Body in Mind: How Gestures Empower Foreign Language Learning. *Mind, Brain, and Education*, *5*(4), 196–211.

Macedonia, M., Müller, K., & Friederici, A. D. (2011). The impact of iconic gestures on foreign language word learning and its neural substrate. *Human Brain Mapping*, *32*(6), 982–998.

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris*, *102*(1), 59–70.

Mahon, B. Z., & Caramazza, A. (2009). Concepts and Categories: A Cognitive Neuropsychological Perspective. *Annual Review of Psychology*, *60*, 27–51.

Mayer, K. M., Yildiz, I. B., Macedonia, M., & von Kriegstein, K. (2015). Visual and motor cortices differentially support the translation of foreign language words. *Current Biology: CB*, *25*(4), 530–535.

McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: Minimal instruction produces rapid change. *Nature Neuroscience*, *7*(7), 703–

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought* (pp. xi, 416). University of Chicago Press.

Mestres-Missé, A., Rodriguez-Fornells, A., & Münte, T. F. (2007). Watching the Brain during Meaning Acquisition. *Cerebral Cortex*, *17*(8), 1858–1866.

Mognon, A., Jovicich, J., Bruzzone, L., & Buiatti, M. (2011). ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology*, *48*(2), 229–240.

Moreno, I., de Vega, M., León, I., Bastiaansen, M., Glen Lewis, A., & Magyari, L. (2015). Brain dynamics in the comprehension of action-related language. A time-frequency analysis of mu rhythms. *NeuroImage*, *109*, 50–62.

Moskowitz, G. (1976). The Classroom Interaction of Outstanding Foreign Language Teachers. *Foreign Language Annals*, *9*(2), 135–143.

Neuper, C., & Pfurtscheller, G. (2001). Event-related dynamics of cortical rhythms: Frequency-specific features and functional correlates. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *43*(1), 41–58.

Niccolai, V., Klepp, A., Weissler, H., Hoogenboom, N., Schnitzler, A., & Biermann-Ruben, K. (2014). Grasping hand verbs: Oscillatory beta and alpha correlates of action-word processing. *PloS One*, *9*(9), e108059.

Nishitani, N., Schürmann, M., Amunts, K., & Hari, R. (2005). Broca's region: From action to language. *Physiology (Bethesda, Md.)*, *20*, 60–69.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2010) *FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data* [Research Article]. Computational Intelligence and Neuroscience; Hindawi.

Öttl, B., Dudschig, C., & Kaup, B. (2017). Forming associations between language and sensorimotor traces during novel word learning. *Language and Cognition*, *9*(1), 156–171.

Papesh, M. H. (2015). Just out of reach: On the reliability of the action-sentence compatibility effect. *Journal of Experimental Psychology. General*, *144*(6), e116-141.

Peeters, D. (2019). Virtual reality: A game-changing method for the language sciences. *Psychonomic Bulletin & Review*, *26*(3), 894–900.

Pfurtscheller, G., & Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: Basic principles. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *110*(11), 1842–1857.

Pineda, J. A. (2005). The functional significance of mu rhythms: Translating "seeing" and "hearing" into "doing." *Brain Research. Brain Research Reviews*, *50*(1), 57–68.

Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences*, *22*(2), 253–336.

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews. Neuroscience*, *6*(7), 576–582.

Shams, L., & Seitz, A. R. (2008). Benefits of multisensory learning. *Trends in Cognitive Sciences*, *12*(11), 411–417.

Shtyrov, Y., Nikulin, V. V., & Pulvermüller, F. (2010). Rapid Cortical Plasticity Underlying Novel Word Learning. *Journal of Neuroscience*, *30*(50), 16864–16867.

Skulmowski, A., & Rey, G. D. (2018). Embodied learning: Introducing a taxonomy based on bodily engagement and task integration. *Cognitive Research: Principles and Implications*, *3*.

Soskey, L., Holcomb, P. J., & Midgley, K. J. (2016). Language effects in second-language learners: A longitudinal electrophysiological study of spanish classroom learning. *Brain Research*, *1646*, 44–52.

Sun, R., Merrill, E., & Peterson, T. (2001). From implicit skills to explicit knowledge: A bottom-up model of skill learning. *Cognitive Science*, *25*(2), 203–244.

Tellier, M. (2008). The effect of gestures on second language memorisation by young children. *Gesture*, *8*(2), 219–235.

Thompson-Schill, S. L. (2003). Neuroimaging studies of semantic memory: Inferring "how" from "where." *Neuropsychologia*, *41*(3), 280–292.

Tomasello, R., Garagnani, M., Wennekers, T., & Pulvermüller, F. (2018). A Neurobiologically Constrained Cortex Model of Semantic Grounding With Spiking Neurons and Brain-Like Connectivity. *Frontiers in Computational Neuroscience*, *12*.

Tromp, J., Peeters, D., Meyer, A. S., & Hagoort, P. (2018). The combined use of virtual reality and EEG to study language processing in naturalistic environments. *Behavior Research Methods*, *50*(2), 862–869.

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

Wakefield, E. M., & James, K. H. (2011). Effects of sensori-motor learning on melody processing across development. *Cognition, Brain, Behavior : An Interdisciplinary Journal*, *15*(4), 505–534.

Wilson, A. D., & Golonka, S. (2013). Embodied Cognition is Not What you Think it is. *Frontiers in Psychology*, *4*.

Yum, Y. N., Midgley, K. J., Holcomb, P. J., & Grainger, J. (2014). An ERP study on initial second language vocabulary learning. *Psychophysiology*, *51*(4), 364–373.

Zappa, A., Bolger, D., Pergandi, J.-M., Mallet, P., Dubarry, A.-S., Mestre, D., & Frenck-Mestre, C. (2019). Motor resonance during linguistic processing as shown by EEG in a naturalistic VR environment. *Brain and Cognition*, *134*, 44–57.

Chapter 7

General Discussion

7.1 Overview

Theories of embodied semantics attribute the body with an essential role in language processing due to the grounding of language in sensory-motor experience (Kiefer & Pulvermüller, 2011; Pulvermüller, 2018). These theories are closely linked to theories of grounded cognition, which credit not only the body but also the surrounding environment with an important role in cognitive processes (Barsalou, 2008). Our studies aimed to add to the literature on embodied semantics by using electrophysiological measures to better define the role of motor processes in semantic representations and language learning. We asked whether, and, if so, how the premotor and motor cortex contribute to the representation of action concepts using cortical measures that can reveal functional connections between semantic and motor processes. A combination of traditional set-ups and virtual reality were employed to explore the role of simulation in both word learning and word processing, using behavioral and electrophysiological measures. We were especially interested in combining cortical measures with ecologically valid methodologies that take into account the importance of the surrounding environment and the body's relationship to it for cognition at large, and specifically language processing.

In the study presented in chapter 2, we taught novice learners a lexicon through interactive computerized games. The study investigated the influence of the L1 on L2 word learning by manipulating L1/L2 grammatical gender congruency across learned linguistic labels. Behavioral and electrophysiological results revealed rapid L2 word learning via games. Post-training, ERP results showed an N400 effect for mismatched audiovisual pairs compared to match trials, but only for words that shared gender across learners' L1 and L2. These results indicate an influence of L1/L2 gender congruency on L2 learning, as well as the effectiveness of interactive games for language learning. In our second study (chapter 3), we combined an Action-sentence compatibility effect (ACE) paradigm and EEG to investigate motor-semantic interactions. We found no effect of language-motor compatibility on motor preparation. On the other hand, congruent motor and semantic trials showed interference on semantic processing, as indexed by an N400 effect for compatible pairs compared to incompatible pairs. This pattern of results is in line with the idea that we perform simulations in order to understand action language (Barsalou, 1999). In our third study (chapter 4), a combined EEG-VR set-up using a Go-Nogo task showed motor resonance during verb processing and prior to movement. This was true for both conditions, whether an action was required or not, bolstering the claim that sensory-motor processes are involved in conceptual representation of linguistic information. Further, we found greater mu desynchronization on Go trials, compared to Nogo trials, suggesting that movement preparation can interact with semantic processing. Together our results support the involvement of motor processes in language processing but do not as such prove causality. Our final projected study is a registered report that seeks to investigate the relationship between motor resonance and linguistic representation using a learning paradigm that manipulates the specificity of action during L2 verb learning. If verbs learned with compatible actions produce greater motor resonance and are better encoded than those learned in the control condition, in which a non-specific action is performed, this would indicate that motor activation can play an essential role in language learning and language processing. The studies described used cortical measures to examine linguistic and motor-semantic effects linked to native language processing and L2 learning. We examined ERPs and time-frequency measures in combination with interactive computer-games as well as VR environments to observe language processing in contextually rich and immersive environments, in line with embodied semantics. The body of results presented ensues from a series of innovative methodologies that succeeded in capturing motor-semantic interactions during language processing.

7.2 Main results

The first study presented in this dissertation (chapter 2) differs from the others as it did not directly measure motor activation during language processing. It did, however, use computerized games as a first step to exploring the effects of interaction and action observation during word encoding. Generally speaking, immersion and interaction have long been considered important for L2 word learning (Freed et al., 2004; Young et al., 2012). Both online and offline interactive video and computer-games have proven conducive to language acquisition (Sylvén & Sundqvist, 2012). Studies that have assessed the benefit of digital games for L2 learning have correlated game-playing to improvements in L2 vocabulary, classroom participation and verbal fluency (Chik, 2014; Kuppens, 2010; Peterson, 2010; Ranalli, 2008; Sundqvist, 2009; Sylvén & Sundqvist, 2012; Young et al., 2012). These effects have been linked to the multimodality of the information received during game-playing, which allows learners to associate new labels to concepts via auditory and visual stimuli (Gee, 2012). On the other hand, very few studies have provided cortical measures linked to L2 learning via computer-games (O'Neil, Lagarrigue, Newman & Frenck-Mestre, 2016). Study 1 (chapter 2) sought to fill this gap by teaching completely novice learners a new vocabulary in an L2 and assessing lexical integration pre and post-training both behaviorally and through EEG using an audio-visual match-mismatch task.

This study also examined the effect of cross-language gender congruency on early L2 lexical acquisition in a natural language. It has been shown that experienced bilinguals whose two languages both have grammatical gender are sensitive to gender congruency across languages (for a review see Sá-Leite et al., 2019). The jury is still out, however, as to why cross-language gender effects occur. Studies on cross-language gender effects have shown inconsistent results, depending on the task employed. Costa and colleagues argued that although semantic representations are shared across languages, grammatical features, such as gender, are an inherent property of that entry and thus not shared across languages (Costa et al., 2003). In addition, the majority of studies on this matter have examined experienced bilinguals. In our study we chose to investigate gender congruency effects from the very first stages of L2 exposure. Using a pre and post-training testing design, we taught participants new linguistic labels for known concepts in an L2. We manipulated grammatical gender congruency between the L2 words and their equivalents in the L1. Participants performed a match-mismatch image-auditory word task pre and post-training while behavioral accuracy and EEG were recorded. Pre-training, no differences were found between conditions, either behaviorally or in ERP results. Post-training, behavioral results showed very high accuracy for both conditions. Importantly, EEG results revealed a large N400 effect for mismatch trials compared to match trials only for linguistic labels that had the same gender across the L1 and the L2. Nouns that had the opposite gender across languages did not produce an N400 effect as a function of the pairing between the auditory word and the visual stimulus.

Our results indicate that speakers of gendered languages automatically activate grammatical gender (Boutonnet et al.,2012; Dahan et al., 2000; Lew-Williams & Fernald, 2007) and that gender incongruency across languages can result in inhibition (Morales et al., 2016; Rodríguez-Fornells & Münte, 2016). This could be interpreted as showing that grammatical features are shared across languages, not stored as an independent feature at the lemma level (cf. Levelt et al., 1999), in line with the interactive activation model of processing (Dell, 1986). To our knowledge, ours is the first study to measure the effect of cross-language gender congruency during the early stages of L2 lexical acquisition in a natural language.

Studies 2 and 3 (chapters 3 and 4) directly investigated the role of motor activation during language processing. Our intent was to deepen our understanding of the nature of linguistic representations, the subject of a longstanding debate. On the one hand, traditional cognitive psychology holds that our linguistic repertoire is made up of abstract, amodal symbols (Fodor, 1980; Hummel, 2010). Semantic network, distributional and feature-based models all stipulate that linguistic symbols are defined by other linguistic symbols (Collins & Quillian, 1969; Landauer & Dumais, 1997; McRae, Cree, Seidenberg & McNorgan, 2005). On the other hand, proponents of embodied semantics claim that linguistic representations are grounded in everyday sensory motor experiences (Barsalou, 1999, 2003; Hauk & Tschentscher, 2013). Action language is often used to test theories of embodied semantics given the evident link between the movements it describes and motor processes (Ibáñez & García, 2018; Klepp et al., 2019; Pulvermüller, 2002).

In our second study (chapter 3), we manipulated motor and semantic compatibility using an Action-sentence compatibility effect (ACE) paradigm and EEG to investigate motor and semantic processes and their influence on one another. Although the ACE paradigm has been widely used behaviorally (Diefenbach et al., 2013; Dudschig et al., 2014; Glenberg & Kaschak, 2002; Zwaan & Taylor, 2006), to our knowledge only one other study has used electrophysiological measures to investigate this effect (Aravena et al., 2010). Importantly, the direction and presence of the effect has been shown to vary as a function of timing, rendering behavioral measures insufficient when it comes to finding causal correlations between motor preparation and semantic processing (Ibáñez & García, 2018). We measured brain activity as participants processed auditory action sentences and simultaneously prepared either a compatible or incompatible movement (away from or towards the body).

Our results confirmed an influence of motor processes on the semantic processing of action sentences, as shown by variations in the N400 response. Contrary to previous studies using EEG to measure the neurophysiological correlates of the ACE (Aravena et al., 2010), compatibility between action language and motor planning produced an inhibitory effect on semantic processing, as indexed by a greater negative deflection in the N400, possibly due to motor and semantic processes competing for neural resources. These results suggest that participants simulated the action described in the sentence, which then competed with the motor preparation necessary for performing a compatible action, indicating shared neural substrates for action and semantic processing. This is in line with the hypothesis that simulations can be involved in understanding action language (Barsalou, 1999, 2020). Our results are also in line with behavioral studies showing that movement can be inhibited when it is required shortly after processing action verbs (0-400ms for verbs presented in isolation and up to 1000 ms for sentences) (Boulenger et al., 2006; Buccino et al., 2005).

In contrast to our findings for semantic processing, our ERP results showed no modulations as a function of action sentence compatibility in components associated with motor preparation (MP from -50 to 90 msec and RAP from 200 to 300 msec). Once again, this result does not corroborate that reported by Aravena and colleagues (2010). In their ACE experiment, compatible trials elicited larger MP amplitudes. This was interpreted as semantic priming of motor preparation, an argument considered crucial to the interaction between linguistic and motor processing. However, the difference across studies may be attributed to task demands. In Aravena et al.'s study (2010), participants only had to indicate when they understood a sentence, using the same repetitive hand movement. As in traditional ACE studies, response mapping therefore occurred early in the sentence (Borreggine & Kaschak, 2006). In our study, the direction of the described action could only be understood at the very end of the sentence; therefore response mapping likely occurred simultaneously to semantic processing.

According to the Hand-Action Network Dynamic Language Embodiment (HANDLE) model (García & Ibáñez, 2016), when we access action language, neural networks that support motor action and processes are automatically activated. The model claims that hand-related expressions (HREs) produce, first, supra and later, sub-threshold activity levels. When hand-motor processes overlap or take place within 400 msec of processing an HRE, the two concurrent processes compete for neural resources. Activity is raised above threshold for one process and the other process has difficulty accessing the same resource, causing interference. This is generally indexed in either a slowing down of motor responses or greater difficulty in semantic processing. In our study, response mapping likely co-occurred with participants processing the action verb. In line with García and Ibáñez's HANDLE model (García & Ibáñez, 2016), these results could be interpreted as indicating that facilitation for motor preparation did not occur because the verb processing did not precede motor response by a long enough period (over 400 msec) to produce sub-threshold activity. Instead, the activity produced from action verb processing during motor preparation was supra threshold, which inhibited semantic access, as indexed by our N400 effect for compatible trials.

In addition to carrying out traditional ERP analyses, we conducted a series of time-frequency analyses to further investigate the role of motor activation during sentential processing. We found mu and beta ERD, or motor activation, specifically in sites that have been associated with motor areas, signaling that processing action-related sentences indeed incurred simultaneous activation in motor regions. The level of desynchronization did not vary, however, as a function of action language compatibility. We therefore concluded that while motor activation affected linguistic processing, semantic processing did not influence motor preparation. Behaviorally, we did not see evidence of an Action-sentence compatibility effect in either direction. Several studies have failed to find a behavioral ACE (cf. Papesh, 2015) and the literature shows that the direction of and the very occurrence of the effect is contingent on factors such as timing, perspective, linguistic context, movement type and task (Alemanno et al., 2012; Aravena et al., 2012, Aravena, 2014; Boulenger et al., 2006; Papeo, Vallesi, Isaja & Rumiati, 2009; Shebani & Pulvermüller, 2018; Sato et al., 2008). The combination of ERP and time-frequency analyses allowed us to obtain a more complete view on motor-semantic interactions compared to other studies in this area (Aravena et al., 2010). Importantly, we showed that even in the absence of behavioral data, i.e. no clear ACE in either direction, an interplay between action and meaning can be seen in brain measures, emphasizing the relevance of analyzing the neural correlates of the ACE. Our study adds to the literature that teases apart facilitatory and inhibitory ACE effects in the search for common mechanisms that may subserve motor and linguistic processes.

Our third study, presented in chapter 4, took the investigation of motor activation during language processing further by adopting an ecologically valid virtual reality environment. We developed a novel experimental paradigm using EEG in a fully immersive Cave automatic virtual environment (CAVE) to measure cortical motor activation during language processing. To our knowledge, this is the first study to combine an immersive CAVE and EEG for observing motor-semantic interactions. Compared to previous studies that examined the neural signature of these interactions (Alemanno et al., 2012; Fargier et al., 2012; Moreno et al., 2015; van Elk et al., 2010), ours used a much richer and closer-to-life environment. Not only were participants visually immersed in a 3D virtual environment; crucially, they were able to manipulate virtual objects using naturalistic actions (e.g. *throw*), adding to their sense of agency (Johnson-Glenberg, 2018). This is especially relevant in the context of embodied semantics, which considers an agent's surrounding environment to be an essential part of cognition (Shapiro, 2011).

Participants performed a Go-Nogo task in which they had to either perform or refrain from performing an action described by an auditory verb (e.g. "jette" [throw]). ERPs and oscillatory activity during action verb processing and prior to movement were analyzed. We found clear ERP components associated with lexical access of the auditory verbs (N1, P2, N400), but no differences in the ERP waveforms as a function of condition. As outlined in the introduction, ERPs are limited when it comes to measuring motor activation as they can only reveal phase-locked activity. ERSP, in contrast, takes both phase-locked and unlocked cortical activity into account. Time-frequency analyses revealed both mu and beta desynchronization during verb processing for both Go and Nogo trials. Further, mu desynchronization was enhanced for Go trials, suggesting that motor processes can be involved in language processing (Alemanno et al., 2012; Moreno et al., 2015; van Elk et al., 2010). These results provide compelling evidence in a naturalistic setting of how motor and linguistic processes interact at the cortical level. In line with embodied semantics, our study fully acknowledged the influence of the physical environment surrounding the participants' mental processes; we believe this renders our results more applicable to real life than what is generally found in controlled neurolinguistic experiments.

Both studies 2 and 3 (chapters 3 and 4) evidenced an influence of motor processes on semantic processing. These studies used paradigms adapted from experiments that measured the impact of the co-occurrence of motor and semantic processes on each other, behaviorally, with the addition of cortical measures. Specifically, the ACE (chapter 3) and Go-Nogo (chapter 4) tasks were used to gain insight as to the precise timing of motor activation in relation to semantic processing. Both studies strongly suggest the existence of shared neural substrates for action and the semantic processing of action language. Given the constraints of the tasks used, one caveat is that neither study included an abstract language control condition. This inclusion would have allowed for a direct comparison between motor activation during action language versus during abstract language processing (Gonzalez, Vega & León, 2013; Moreno et al., 2015; Sakreida et al., 2013). Had motor resonance been greater for action language, for example, we would have more easily established a direct link between motor activation and semantic meaning of the features of the word being processed. As they stand, neither study can claim to prove that the motor activations found are an essential part of language processing. As pointed out by Fargier (2016), isolating the functionality of motor activation for language processing is extremely difficult when examining a well-established native language. One way around this, he claims, is to use learning paradigms that train learners to associate motor content to novel linguistic labels, and neural measures to examine the extent to which language representations are embodied.

The projected study in chapter 6, submitted as a registered report, is the culmination of our first three studies. Within the framework of embodied semantics, the report proposes to explore the role of motor activation during word learning. While study 1 (chapter 2) used interaction and action observation via computer-games to enhance learning, embodied learning has mostly been associated with "self-performed" or "self-generated" actions (James & Bose, 2011; James & Swain, 2011; Johnson-Glenberg, 2017, 2018; Johnson-Glenberg & Megowan-Romanowicz, 2017). Furthermore, learning can only be considered truly embodied in the presence of a direct one-to-one correspondence between the lexical item being learned and the action being performed (Johnson-Glenberg & Megowan-Romanowicz, 2017; Skulmowski & Rey, 2018). With this in mind, in the registered report in chapter 6, we aim to teach participants L2 action verbs as they perform corresponding actions. We will apply a similar pre/post-training testing design as was used in Study 1. Both pre and post-training, behavioral and ERP measures will be recorded during a match-mismatch task, involving auditory action verbs and video clips of actions. In addition, we will conduct time-frequency analyses during passive verb listening. Having found mu and beta desynchronization both during single verb processing and during action-sentence processing, in studies 2 and 3 (chapters 3 and 4), we decided to further investigate the origin of this activation.

The Hebbian theory of associative learning claims that acquiring lexical items along with physical movement creates neural assemblies that connect action planning and execution to semantic processing (Hebb, 1949; Pulvermüller, 1999, 2005; Tomasello et al., 2018). Another explanation as to why these two processes could form common representations is Zwaan and Madden's (2005) theory of experiential traces, according to which linguistic labels become associated with "experiential traces" formed as we interact with the environment. The registered report in chapter 5 aims to directly explore whether learning action verbs in an L2 using specific movements that map onto action verbs produces greater motor activation post-training compared to a control condition that only involves pointing. Further, we ask whether increased motor activation during verb processing post-training correlates with improved learning. Fargier et al. (2012) found no behavioral differences between learning novel words by associating them to specific hand actions compared to associating them to abstract animations. Their results did, however, show greater motor activation, in the form of mu desynchronization, for words learned with physical actions, but only on day 1 of learning. On day 2, they found activity in a "convergence zone" between motor and language structures instead. Our registered report is designed to build on these results using a more ecologically valid approach, both linguistically, as we used a natural language as opposed to pseudo-words, and as concerns the learning environment. We advanced in our use of virtual reality, which began in Study 3 (chapter 4), to develop an interactive, realistic 3D environment in which participants can perform naturalistic movements and manipulate a virtual object.

An extensive amount of experimental evidence has shown that motor structures are recruited during language processing. However, the debate as to whether these structures contribute to or are necessary for semantic access remains very much alive. The difficulty in resolving this debate lies in showing functionality of motor activation during language processing. As such, learning provides an avenue through which to trace this activation from encoding to processing. Our registered report aims to observe the effects of associating induced motor activation during language encoding specifically to measure how this affects learning outcome and language processing. More generally, it will serve as a means to gain insight into the functional role of motor structures in language processing (Fargier et al., 2012).

7.3 Our studies and models of embodied semantics

Since its beginnings, the embodiment debate has evolved from simply contradicting computational and abstract views on cognition to a highly nuanced and complex discussion over the exact role of the body and the environment in cognition. Indeed, the term "embodied" can be limited in describing what it encompasses, which is not only the actual body's involvement in cognition, but also sensory-motor systems and the physical and social environments. Hence "grounded cognition" or Barsalou's "situated conceptualization" might better describe the phenomena we were interested in examining (Barsalou, 2008, 2010, 2016). However, a more recent sub-category of embodied cognition has emerged, which pertains directly to language and specifically semantic processing: embodied semantics. For the sake of simplicity, we have chosen to use this term throughout the majority of this dissertation as we believe a great deal of the state of the art related to our work situates itself within this framework. As detailed in the introduction, embodied semantics attempts to elucidate the exact role of motor involvement in conceptual representations, whether as an integral part of semantic representation, or simply reflective of post-lexical imagery processes. Empirical evidence acquired over the last 20 years has established that motor processes are at least partially involved in semantic processes as seen through the influence of meaning on motor processes and the activation of sensory-motor regions during semantic processing (Barsalou, 1999; Coello & Fischer, 2016; Fischer & Zwaan, 2008). Models of embodied semantics therefore posit that sensory-motor mechanisms and contextual cues determine the workings of high-level cognitive functions (Coello & Fischer, 2016; Fischer & Zwaan, 2008; Pulvermüller, 2013). However, models differ when it comes to accounting for how, when and why this is the case. In the following sections we situate our research within influential theories of embodied semantics (Barsalou, 1999; Pulvermüller, 1999, 2013).

The observation that many category-specific activations occur in cortical areas that overlap or are adjacent to sensory-motor areas (i.e. object words activate ventral visual areas and action words activate the dorsal motor cortex) has lead to a neuromechanistic approach to semantic processing, namely that shared mechanisms exist between semantic and action and perception processes (Kiefer & Pulvermüller, 2012; Pulvermüller, 2005). Underlying this framework, is the "correlational learning principle" according to which the co-occurrence of action-perception and meaning results in the common firing of neurons that form "embodied referential semantic circuits" to support meaning representation (Hebb, 1949; Pulvermüller, 2013). Hence, the same neural networks that support the acquisition of meaning are thought to subserve representation and processing once semantic information has been encoded (Tschentscher, 2017). As outlined in the introduction, a great deal of neuroimaging evidence points to the recruitment of motor structures during semantic processing. This does not, however, necessarily indicate that this recruitment is functional or necessary for language understanding (Mahon & Caramazza, 2008). Studies showing early involvement of motor structures either through interference or facilitation, i.e. within 200-300 msec of language input, as opposed to later, possibly post-lexical effects, are more convincing (Boulenger et al., 2009; Mollo et al., 2016). Notwithstanding, the earliness of the effect does not suffice to indicate functionality, given the possibility of spreading activation (Mahon & Caramazza, 2008; Tschentscher, 2017).

In our ACE study (chapter 3), we did not succeed in finding such early effects. Rather an inhibitory effect for action-meaning compatible trials was only seen the N400 window, between 300-600 msec post verb offset. One explanation may lie in the complexity of the language being processed. Indeed the sentences used as stimuli included two clauses and the action performed in the sentence was only unveiled by the verb at the very end of the sentence. Other studies have also shown later, end-of-sentence motor activation during language processing (Moreno et al., 2015). Aravena and colleagues found greater motor preparation at the onset of the sentence-final verb and their semantic effects took place in roughly the same time window as ours (Aravena et al., 2010). Given the lateness of our effects, we cannot argue that the mere inhibition of semantic processes resulting from motor response evidences a functional link between the two. However, the discrepancies between the timing of our effects and the early effects shown in single verb studies underline the need for situating action language within richer linguistic contexts that come close to reflecting real-life language use. In addition, the differences in movement complexity (stretching or recoiling one's arm as opposed to repetitive button pressing) could account for the discrepancies between ours and Aravena and colleagues' results (Shebani & Pulvermüller, 2018). Again, future studies would benefit from using more ecological actions in order to further this line of research.

The results from our body of work showing motor-semantic interactions (chapters 3 and 4) could also be interpreted within the framework of simulations for action language comprehension. This idea originates in one of the earliest and most influential models that attempt to account for embodied and grounded language processes, Barsalou's theory of perceptual symbols (1999). The theory claims that multimodal conceptual representations encompass the modal processes involved in experience. These "situated conceptualizations" are combined through a mechanism, often referred to as "simulation", that underlies language comprehension (Barsalou, 2016). Within this framework, simulations are essential for cognition given their involvement in both online processing and off-line access to concepts, linking new information to perceptual and motor processes experienced in the past (Barsalou, 1999, 2008; Coello & Fischer, 2016; Jeannerod, 2006). Cognition, generally, and language processing specifically, are therefore thought to be grounded in the physical body and situated in past experience (Coello & Fischer, 2016). This model does not rule out the existence of conceptual representations; on the contrary, it considers that concepts are essential for "representing and processing non-present situations" and hence for cognition (Barsalou, 2016). The theory holds that representational processes are not limited to simulations and that other representational mechanisms are necessary for language processing, without specifying which ones. Not unlike the idea of convergence zones, Barsalou admits the possibility of abstract representations that could take place in association areas (Barsalou, 2008; Barsalou, 2016; Simmons & Barsalou, 2003). Importantly, he has also recently pointed out that, although his theory is consistent with much of the evidence coming out of embodied semantic studies, very little of this evidence directly validates the existence of situated conceptualizations and their role in cognition (Barsalou, 2016). Our results showing that motor processes interfered with semantic processing (chapter 3) are in line with the idea that simulations support language processing. Once again, our results do not provide direct evidence of this phenomenon but add to the body of evidence that supports it (Aravena et al., 2010; Boulenger et al., 2009; Glenberg &‘ Kaschak, 2002; van Elk et al., 2010; Zwaan & Taylor, 2006).

Pulvermüller's neurobiological view of embodied and grounded concepts originates in the Hebbian principle of distributed networks co-activated during initial encoding supporting meaning (Hebb, 1949; Pulvermüller, 2013). Recently, hybrid approaches of embodiment, sometimes called "weak embodiment", have attempted to account for both evidence of motor involvement in language processing and its non-essential role (Tirado et al., 2018; Meteyard, Cuadrado, Bahrami & Vigliocco, 2012). Proponents of these models often rely on integration areas, such as convergence zones, as a middle-ground solution between abstract and modal representations (Galetzka, 2017). Convergence zones, originally proposed by Damasio (1989) are thought to be located between language and sensory-motor areas. Their function is to relay signals from different modality-specific regions, allowing them to cooperate during simulation of sensory-motor events (Damasio & Tranel, 1993; Fargier, 2013; Meyer & Damasio, 2009; Simmons & Barsalou, 2003). In this view, semantic information is not stored in the primary modal cortices but in adjacent association areas (Fargier et al., 2012; Simmons & Barsalou, 2003). It has further been suggested that semantic processing only partially recruits sensory-motor structures involved in perception and motor experiences because a full simulation, or activation of structures involved in sensory-motor processes, would compete with motor preparation (Willems, Toni, Hagoort & Casasanto, 2009). Along similar lines, Fargier and colleagues (2012) observed mu desynchronization in central-parietal electrodes, and interpreted this as indicative of motor activation resulting from participants linking novel words and manual actions. However, as these words became better integrated, mu suppression was observed over fronto-central areas, associated with convergence zones. The authors concluded that the motor activation that results from language processing indicates that motor representations are underspecified as opposed to explicit and that language representations are hence not dependent on sensory-motor activations that take place during language encoding. They maintain, however, that concepts are grounded in perception and action (Fargier et al., 2012).

7.4 Conclusion

The line of investigation proposed by embodied, grounded and situated cognition is an exciting one. It points to a path away from computational models and toward a more holistic approach to cognition. A considerable amount of literature has explored motor-semantic interactions within embodied frameworks, both behaviorally and through cortical measures. These two approaches are, of course, complementary and necessary given the complexity of the investigations at hand. Indeed, the very questions we ask — Are “representations” abstract? Do they even exist? Is “simulation” involved in or necessary for language processing? What role does motor simulation play in “storing” or “accessing” language? — ‘sway us in an initial direction that can be misleading. These inquiries often lead to contradictory results given the crucial role of a) the temporal relationship between motor activation and language processing, b) the various measures employed and c) the different paradigms and environments used. The embodied semantics literature has come a long way in unveiling subtle differences between semantic-to-motor and motor-to-semantic effects due to contingencies. However, it also seems to often “hit a wall” when it comes to providing evidence for embodied theories, mainly due to the limits of examining embodied processes in a disembodied manner. A solution lies in designing new, inventive paradigms that bring the body, action and the surrounding environment within the limits of experimental control. The studies in this dissertation add a valuable piece to the motor-semantic interaction puzzle by virtue of their combination of brain measures and interactive, ecologically valid experimental paradigms. Whether we believe that language is symbolic, embodied or somewhere in between, real-life language processing is undeniably dynamic and contextual. A methodological evolution is called for and our studies have helped to pave the path toward approaches and paradigms that will better elucidate the true nature of language processing.

7.5 Perspectives

An evident aim for future research is to fully exploit the developed VR/EEG paradigms as a multimodal experimental approach. Although we succeeded in combining EEG and a CAVE in our third study, the task performed represents but a first step in using this novel methodology to test embodied theories. Grounded cognition, particularly, binds social and physical contexts to cognition (Barsalou, 2010). It would be of interest to manipulate the actual environment and context in order to examine how language processing and L2 retention are affected by interactive and situated experiences. For example, environments could be manipulated to create physical contexts that are more or less related to presented language so as to measure the outcome on learning and processing meaning (e.g. teaching a food-related lexicon in an L2 to novice learners in a virtual grocery store as compared to on a beach). Furthermore, avatars or 3D representations of agents could be manipulated to explore the role of social interactions on language processing and learning. Once again, EEG could be employed either in the virtual reality environment or pre and post-training, in the case of learning paradigms. Finally, it would be beneficial for the L2 learning and the embodied literature to develop longitudinal embodied L2 learning paradigms in order to provide a more in-depth account of how the body and the environment can support learning in the long run. Generally speaking, coupling EEG and VR environments is of interest to the VR community at large, as demonstrated by both theoretical (Peeters, 2019; Tromp, Peeters, Meyer & Hagoort, 2018) and methodological (chapter 5; Hertweck et al., 2019) publications.

References

Alemanno, F., Houdayer, E., Cursi, M., Velikova, S., Tettamanti, M., Comi, G., Cappa, S. F., & Leocani, L. (2012). Action-related semantic content and negation polarity modulate motor areas during sentence reading: An event-related desynchronization study. *Brain Research*, *1484*, 39–49.

Amsel, B. D., Urbach, T. P., & Kutas, M. (2013). Alive and grasping: Stable and rapid semantic access to an object category but not object graspability. *NeuroImage*, *77*, 1–13.

Aravena, P. (2014). *Dynamics of language induced cortical motor activity: —Determining the linguistic contexts that trigger motor activation during lexical semantic processing*. [Doctoral dissertation, Université Claude Bernard Lyon 2]

Aravena, P., Courson, M., Frak, V., Cheylus, A., Paulignan, Y., Deprez, V., & Nazir, T. (2014). Action relevance in linguistic context drives word-induced motor activity. *Frontiers in Human Neuroscience*, *8*.

Aravena, P., Delevoye-Turrell, Y., Deprez, V., Cheylus, A., Paulignan, Y., Frak, V., & Nazir, T. (2012). Grip Force Reveals the Context Sensitivity of Language-Induced Motor Activity during "Action Words" Processing: Evidence from Sentential Negation. *PLOS ONE*, *7*(12), e50287.

Aravena, P., Hurtado, E., Riveros, R., Cardona, J. F., Manes, F., & Ibáñez, A. (2010). Applauding with closed hands: Neural signature of action-sentence compatibility effects. *PloS One*, *5*(7), e11751.

Ardal, S., Donald, M. W., Meuter, R., Muldrew, S., & Luce, M. (1990). Brain responses to semantic incongruity in bilinguals. *Brain and Language*, *39*(2), 187–205.

Atkinson, D. (2002). Toward a Sociocognitive Approach to Second Language Acquisition. *The Modern Language Journal*, *86*(4), 525–545. JSTOR.

Atkinson, D. (2010). Extended, Embodied Cognition and Second Language Acquisition. *Applied Linguistics*, *31*(5), 599–622.

Aziz-Zadeh, L., & Damasio, A. (2008). Embodied semantics for actions: Findings from functional brain imaging. *Journal of Physiology, Paris*, *102*(1–3), 35–39.

Barber, H. A., & Kutas, M. (2007). Interplay between computational models and cognitive electrophysiology in visual word recognition. *Brain Research Reviews*, *53*(1), 98–123.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(4), 577–660.

Barsalou, L. W. (2003). Situated simulation in the human conceptual system. *Language and Cognitive Processes*, *18*(5–6), 513–562.

Barsalou, L. W. (2008). Grounded Cognition. *Annual Review of Psychology*, *59*(1), 617–645.

Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1281–1289.

Barsalou, L. W. (2010). Grounded Cognition: Past, Present, and Future. *Topics in Cognitive Science*, *2*(4), 716–724.

Barsalou, L. W. (2016). On Staying Grounded and Avoiding Quixotic Dead Ends. *Psychonomic Bulletin & Review*, *23*(4), 1122–1142.

Barsalou, L. W., Dutriaux, L., & Scheepers, C. (2018). Moving beyond the distinction between concrete and abstract concepts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1752), 20170144.

Barsalou, L. W. (2020). Challenges and Opportunities for Grounding Cognition. *Journal of Cognition*, *3*(1), 31.

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

Beer, R. D. (1990). *Intelligence as Adaptive Behaviour: An Experiment in Computational Neuroethology* (First Edition). Academic Press.

Beer, R. D. (2003). The Dynamics of Active Categorical Perception in an Evolved Model Agent. *Adaptive Behavior*, *11*(4), 209–243.

Binder, J. R., & Desai, R. H. (2011). The Neurobiology of Semantic Memory. *Trends in Cognitive Sciences*, *15*(11), 527–536.

Black, J. B., Segal, A., Vitale, J., & Fadjo, C. (2012). Embodied cognition and enhancing learning and motivation. In *Theoretical Foundations of Learning Environments* (D. Jonassen and S. Land, p. 2015).

Bohil, C. J., Alicea, B., & Biocca, F. A. (2011). Virtual reality in neuroscience research and therapy. *Nature Reviews Neuroscience*, *12*(12), 752–762.

Borovsky, A., Elman, J. L., & Kutas, M. (2012). Once is Enough: N400 Indexes Semantic Integration of Novel Word Meanings from a Single Exposure in Context. *Language Learning and Development: The Official Journal of the Society for Language Development*, *8*(3), 278–302.

Borreggine, K. L., & Kaschak, M. P. (2006). The Action–Sentence Compatibility Effect: It's All in the Timing. *Cognitive Science*, *30*(6), 1097–1112.

Boulenger, V., Hauk, O., & Pulvermüller, F. (2009). Grasping ideas with the motor system: Semantic somatotopy in idiom comprehension. *Cerebral Cortex (New York, N.Y.: 1991)*, *19*(8), 1905–1914.

Boulenger, V., Roy, A., Paulignan, Y., Deprez, V., Jeannerod, M., & Nazir, T. (2006). Cross-talk between language processes and overt motor behavior in the first 200 msec of processing. *Journal of Cognitive Neuroscience*.

Boutonnet, B., Athanasopoulos, P., & Thierry, G. (2012). Unconscious effects of grammatical gender during object categorisation. *Brain Research*, *1479*, 72–79.

Bowden, H. W., Steinhauer, K., Sanz, C., & Ullman, M. T. (2013). Native-like brain processing of syntax can be attained by university foreign language learners. *Neuropsychologia*, *51*(13), 2492–2511.

Broaders, S. C., Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2007). Making children gesture brings out implicit knowledge and leads to learning. *Journal of Experimental Psychology: General*, *136*(4), 539–550.

Brown, C. M., & Hagoort, P. (2000). On the electrophysiology of language comprehension: Implications for the human language system. *237*.

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

Burdea, G. C., & Coiffet, P. (2003). *Virtual Reality Technology*. John Wiley & Sons.

Caetano, G., Jousmäki, V., & Hari, R. (2007). Actor's and observer's primary motor cortices stabilize similarly after seen or heard motor actions. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, *104*(21), 9058–9062.

Carroll, J. B. (1993). *Human Cognitive Abilities: A Survey of Factor-Analytic Studies*. Cambridge University Press.

Chatterjee, A. (2010). Disembodying cognition. *Language and Cognition*, *2*(1), 79–116.

Chemero, T. (n.d.). *A. Chemero, "Radical embodied cognitive science", for "Unifying Psychology", special issue of Review of General Psychology, 17, 145-150, 2013.*

Chik, A. (2014). Digital gaming and language learning: Autonomy and community. *Language Learning and Technology*, *18*(2), 85–100.

Childers, J. B., & Tomasello, M. (2002). Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Developmental Psychology*, *38*(6), 967–978Chu, M., & Hagoort, P. (2014). Synchronization of speech and gesture: Evidence for interaction in action. *Journal of Experimental Psychology: General*, *143*(4), 1726–1741.

Chu, M., & Hagoort, P. (2014). Synchronization of speech and gesture: Evidence for interaction in action. *Journal of Experimental Psychology: General*, *143*(4), 1726–1741.

Clark, A. (1997). *Being There: Putting Brain, Body and World Together Again*. MIT Press.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204.

Coello, Y., & Bartolo, A. (2013). *Language and Action in Cognitive Neuroscience*. Psychology Press.

Coello, Y., & Fischer, M. (2016). *Foundations of embodied cognition, Perceptual and Emotional Embodiment.* Psychology Press.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407–428.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning & Verbal Behavior*, *8*(2), 240–247.

Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition*, *106*(2), 1047–1058.

Cook, V., & McDonough, S. (n.d.). *First and second language learning*. Retrieved October 7, 2020, from

Costa, A., Kovacic, D., Franck, J., & Caramazza, A. (2003). On the autonomy of the grammatical gender systems of the two languages of a bilingual. *Bilingualism: Language and Cognition*, *6*(3), 181–200.

Cummings, A., Čeponienė, R., Dick, F., Saygin, A. P., & Townsend, J. (2008). A developmental ERP study of verbal and non-verbal semantic processing. *Brain Research*, *1208*, 137–149.

Dahan, D., Swingley, D., Tanenhaus, M., & Magnuson, J. (2000). Linguistic Gender and Spoken-Word Recognition in French. *Journal of Memory and Language - J MEM LANG*, *42*, 465–480.

Damasio, A R, & Tranel, D. (1993). Nouns and verbs are retrieved with differently distributed neural systems. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(11), 4957–4960.

Damasio, Antonio R. (1989). The brain binds entities and events by multiregional activation from convergence zones. *Neural Computation*, *1*(1), 123–132.

de Nooijer, J. A., van Gog, T., Paas, F., & Zwaan, R. A. (2013). When left is not right: Handedness effects on learning object-manipulation words using pictures with left- or right-handed first-person perspectives. *Psychological Science*, *24*(12), 2515–2521.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283–321.

Diefenbach, C., Rieger, M., Massen, C., & Prinz, W. (2013). Action-Sentence Compatibility: The Role of Action Effects and Timing. *Frontiers in Psychology*, *4*.

Dijkstra, T., & Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(3), 175–197.

Dudschig, C., de la Vega, I., & Kaup, B. (2014). Embodiment and second-language: Automatic activation of motor responses during processing spatially associated L2 words and emotion L2 words in a vertical Stroop paradigm. *Brain and Language*, *132*, 14–21.

Eichert, N., Peeters, D., & Hagoort, P. (2018). Language-driven anticipatory eye movements in virtual reality. *Behavior Research Methods*, *50*(3), 1102–1115.

Elgort, I., Perfetti, C. A., Rickles, B., & Stafura, J. Z. (2015). Contextual learning of L2 word meanings: Second language proficiency modulates behavioural and ERP indicators of learning*. *Language, Cognition and Neuroscience*, *30*(5), 506–528.

Engelkamp, J. (1980). Some studies on the internal structure of propositions*. Psychological Research, 41* (355-371).

Engelkamp, J., & Krumnacker, H. (1980). Image- and motor-processes in the retention of verbal materials. *Zeitschrift Für Experimentelle Und Angewandte Psychologie*, *27*(4), 511–533.

Engelkamp, J., & Zimmer, H. D. (1984). Motor programme information as a separable memory unit. *Psychological Research*, *46*(3), 283–299.

Esopenko, C., Gould, L., Cummine, J., Sarty, G. E., Kuhlmann, N., & Borowsky, R. (2012). A neuroanatomical examination of embodied cognition: Semantic generation to action-related stimuli. *Frontiers in Human Neuroscience*, *6*.

Fargier, R., (2013) *Cerveau et Sens des mots: De l'émergence à la flexibilité des représentations sémantiques dans le cerveau.* [Doctoral dissertation, Université Claude Bernard Lyon 1.]

Fargier, R. (2016). From modal brain structures to word meaning. In *Foundations of embodied cognition: Conceptual and interactive embodiment* (pp. 11–28). Routledge/Taylor & Francis Group.

Fargier, R., Paulignan, Y., Boulenger, V., Monaghan, P., Reboul, A., & Nazir, T. A. (2012). Learning to associate novel words with motor actions: Language-induced motor activity following short training. *Cortex*, *48*(7), 888–899.

Fischer, M. H., & Coello, Y. (2015). *Conceptual and Interactive Embodiment: Foundations of Embodied Cognition Volume 2*. Routledge.

Fischer, M. H., & Zwaan, R. A. (2008). Embodied Language: A Review of the Role of the Motor System in Language Comprehension: *Quarterly Journal of Experimental Psychology*.

Fodor, J. A. (1980). Methodological solipsism considered as a research strategy in cognitive psychology. *Behavioral and Brain Sciences*, *3*(1), 63–109.

Fodor, Jerry A. (1987). Why Paramecia Don't Have Mental Representations. *Midwest Studies In Philosophy*, *10*(1), 3–23.

Fodor, Jerry A. (1994). *The Elm and the Expert: Mentalese and Its Semantics*. MIT Press.

Fodor, Jerry A. (1998). *Concepts: Where Cognitive Science Went Wrong*. Oxford University Press.

Freed, B. F., Segalowitz, N., & Dewey, D. (2004). Context of Learning and Second Language Fluency in French : Comparing Regular Classroom, Study Abroad, and Intensive Domestic Immersion Programs. *Studies in Second Language Acquisition*, *26*(2), 275–301.

Friedrich, M., & Friederici, A. D. (2008). Neurophysiological correlates of online word learning in 14-month-old infants. *Neuroreport*, *19*(18), 1757–1761.

Friedrich, M., Wilhelm, I., Born, J., & Friederici, A. D. (2015). Generalization of word meanings during infant sleep. *Nature Communications*, *6*(1), 6004.

Gainotti, G. (2006). Anatomical functional and cognitive determinants of semantic memory disorders. *Neuroscience and Biobehavioral Reviews*, *30*(5), 577–594.

Galetzka, C. (2017). The Story So Far: How Embodied Cognition Advances Our Understanding of Meaning-Making. *Frontiers in Psychology*, *8*.

Garagnani, M., & Pulvermüller, F. (2016). Conceptual grounding of language in action and perception: A neurocomputational model of the emergence of category specificity and semantic hubs. *European Journal of Neuroscience*, *43*(6), 721–737.

García, A. M., & Ibáñez, A. (2016). A touch with words: Dynamic synergies between manual actions and language. *Neuroscience and Biobehavioral Reviews*, *68*, 59–95.

Gee, J.P., H. (2012). *Foreword. In: Reinders, H., Digital Games in Language Learning and Teaching*. Palgrave Macmillan UK.

Geranmayeh, F., Leech, R., & Wise, R. J. S. (2016). Network dysfunction predicts speech production after left hemisphere stroke. *Neurology*, *86*(14), 1296–1305.

Ghitza, O. (2017). Acoustic-driven delta rhythms as prosodic markers. *Language, Cognition and Neuroscience*, *32*(5), 545–561.

Gianelli, C., Farnè, A., Salemme, R., Jeannerod, M., & Roy, A. C. (2011). The Agent is Right: When Motor Embodied Cognition is Space-Dependent. *PLoS ONE*, *6*(9).

Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience*, *15*(4), 511–517.

Glenberg, A., Havas, D., Becker, R., & Rinck, M. (2005). Grounding language in bodily states: The case for emotion. *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thinking*, 115–128.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, *9*(3), 558–565.

Gonzalez, I., Vega, M., & León, I. (2013). Understanding action language modulates oscillatory mu and beta rhythms in the same way as observing actions. *Brain and Cognition*, *82*, 236–242.

Grandchamp, R., & Delorme, A. (2011). Single-trial normalization for event-related spectral decomposition reduces sensitivity to noisy trials. *Frontiers in Psychology*, *2*, 236.

Grosjean, F., & Byers-Heinlein, K. (2018). *The Listening Bilingual: Speech Perception, Comprehension, and Bilingualism*. John Wiley & Sons.

Hahne, A. (2001). What's Different in Second-Language Processing? Evidence from Event-Related Brain Potentials. *Journal of Psycholinguistic Research*, *30*(3), 251–266.

Hahne, A., & Friederici, A. D. (2002). Differential task effects on semantic and syntactic processes as revealed by ERPs. *Brain Research. Cognitive Brain Research*, *13*(3), 339–356.

Hamm, J. P., Johnson, B. W., & Kirk, I. J. (2002). Comparison of the N300 and N400 ERPs to picture stimuli in congruent and incongruent contexts. *Clinical Neurophysiology*, *113*(8), 1339–1350.

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

Hauk, O. (2016). Only time will tell – why temporal information is essential for our neuroscientific understanding of semantics. *Psychonomic Bulletin & Review*, *23*(4), 1072–1079.

Hauk, O., Giraud, A.-L., & Clarke, A. (2017). Brain oscillations in language comprehension. *Language, Cognition and Neuroscience*, *32*(5), 533–535.

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

Hauk, O., & Tschentscher, N. (2013). The body of evidence: What can neuroscience tell us about embodied semantics? *Frontiers in Psychology*, *4*.

Hebb, D. O. (1949). *The organization of behavior; a neuropsychological theory* (pp. xix, 335). Wiley.

Hertweck, S., Weber, D., Alwanni, H., Unruh, F., Fischbach, M., Latoschik, M., & Ball, T. (2019). *Brain Activity in Virtual Reality: Assessing Signal Quality of High-Resolution EEG While Using Head-Mounted Displays* (p. 971).

Heyselaar, E., Hagoort, P., & Segaert, K. (2017). In dialogue with an avatar, language behavior is identical to dialogue with a human partner. *Behavior Research Methods*, *49*(1), 46–60.

Hobson, H., & Bishop, D. V. M. (2016). Mu suppression – A good measure of the human mirror neuron system? *Cortex*.

Holcomb, P. J. (1993). Semantic priming and stimulus degradation: Implications for the role of the N400 in language processing. *Psychophysiology*, *30*(1), 47–61.

Hummel, J. E. (2010). Symbolic Versus Associative Learning. *Cognitive Science*, *34*(6), 958–965.

Hurley, S. L. (1998). *Consciousness in Action*. Harvard University Press.

Ibáñez, A., & García, A. M. (2018). *Contextual Cognition: The Sensus Communis of a Situated Mind*. Springer International Publishing.

James, K. H., & Bose, P. (2011). Self-generated actions during learning objects and sounds create sensori-motor systems in the developing brain. *Cognition, Brain, Behavior : An Interdisciplinary Journal*, *15*(4), 485–503.

James, K. H., & Swain, S. N. (2011). Only self-generated actions create sensori-motor systems in the developing brain. *Developmental Science*, *14*(4), 673–678.

Jeannerod, M. (2006). *Motor cognition: What actions tell the self* (pp. x, 209). Oxford University Press.

Johnson-Glenberg, M. C. (2017). Embodied Education in Mixed and Mediated Realties. In D. Liu, C. Dede, R. Huang, & J. Richards (Eds.), *Virtual, Augmented, and Mixed Realities in Education* (pp. 193–217). Springer.

Johnson-Glenberg, M. C. (2018). Immersive VR and Education: Embodied Design Principles That Include Gesture and Hand Controls. *Frontiers in Robotics and AI*, *5*.

Johnson-Glenberg, M. C., & Megowan-Romanowicz, C. (2017). Embodied science and mixed reality: How gesture and motion capture affect physics education. *Cognitive Research: Principles and Implications*, *2*(1), 24.

Johnson-Glenberg, M. C., Megowan-Romanowicz, C., Birchfield, D. A., & Savio-Ramos, C. (2016). Effects of Embodied Learning and Digital Platform on the Retention of Physics Content: Centripetal Force. *Frontiers in Psychology*, *7*.

Kiefer, M., & Pulvermüller, F. (2012). Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. *Cortex*, *48*(7), 805–825.

Klepp, A., Dijk, H. van, Niccolai, V., Schnitzler, A., & Biermann-Ruben, K. (2019). Action verb processing specifically modulates motor behaviour and sensorimotor neuronal oscillations. *Scientific Reports*, *9*(1), 1–11.

Knoeferle, P. (2015). Language comprehension in rich non-linguistic contexts: Combining eye-tracking and event-related brain potentials. In R. M. Willems (Ed.), *Cognitive Neuroscience of Natural Language Use* (pp. 77–100). Cambridge University Press.

Koelewijn, T., van Schie, H. T., Bekkering, H., Oostenveld, R., & Jensen, O. (2008). Motor-cortical beta oscillations are modulated by correctness of observed action. *NeuroImage*, *40*(2), 767–775.

Kontra, C., Lyons, D. J., Fischer, S. M., & Beilock, S. L. (2015). Physical Experience Enhances Science Learning: *Psychological Science*.

Kösem, A., & van Wassenhove, V. (2017). Distinct contributions of low- and high-frequency neural oscillations to speech comprehension. *Language, Cognition and Neuroscience*, *32*(5), 536–544.

Kuppens, A. H. (2010). Incidental foreign language acquisition from media exposure. *Learning, Media and Technology*, *35*(1), 65–85.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology*, *62*, 621–647.

Kutas, M., & Hillyard, S. A. (1980). Event-related brain potentials to semantically inappropriate and surprisingly large words. *Biological Psychology*, *11*(2), 99–116.

Ladouce, S., Donaldson, D. I., Dudchenko, P. A., & Ietswaart, M. (2016). Understanding Minds in Real-World Environments: Toward a Mobile Cognition Approach. *Frontiers in Human Neuroscience*, *10*, 694.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*(2), 211–240.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1–38.

Lew-Williams, C., & Fernald, A. (2007). Young Children Learning Spanish Make Rapid Use of Grammatical Gender in Spoken Word Recognition. *Psychological Science*, *18*(3), 193–198.

Litcofsky, K. A., & Van Hell, J. G. (2017). Switching direction affects switching costs: Behavioral, ERP and time-frequency analyses of intra-sentential codeswitching. *Neuropsychologia*, *97*, 112–139.

Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique*. MIT Press.

Macedonia, M., & Klimesch, W. (2014). Long-Term Effects of Gestures on Memory for Foreign Language Words Trained in the Classroom. *Mind, Brain, and Education*, *8*(2), 74–88.

Macedonia, M., & Knösche, T. R. (2011). Body in Mind: How Gestures Empower Foreign Language Learning. *Mind, Brain, and Education*, *5*(4), 196–211.

Machery, E. (2016). The amodal brain and the offloading hypothesis. *Psychonomic Bulletin & Review*, *23*(4), 1090–1095.

MacWhinney, B. (2009). A Unified Model of Language Acquisition. In *Handbook of bilingualism: Psycholinguistic approaches*. Oxford University Press.

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris*, *102*(1), 59–70.

Makeig, S. (1993). Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones. *Electroencephalography & Clinical Neurophysiology*, *86*(4), 283–293.

Makeig, S., Debener, S., Onton, J., & Delorme, A. (2004). Mining event-related brain dynamics. *Trends in Cognitive Sciences*, *8*(5), 204–210.

Matsumoto, M., Okada, E., Ichihara, D., Chiba, K., Toyama, Y., Fujiwara, H., Momoshima, S., Nishiwaki, Y., Hashimoto, T., Inoue, T., Watanabe, M., & Takahata, T. (2010). Prospective ten-year follow-up study comparing patients with whiplash-associated disorders and asymptomatic subjects using magnetic resonance imaging. *Spine*, *35*(18), 1684–1690.

Mayer, K. M., Yildiz, I. B., Macedonia, M., & von Kriegstein, K. (2015). Visual and motor cortices differentially support the translation of foreign language words. *Current Biology: CB*, *25*(4), 530–535.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought* (pp. xi, 416). University of Chicago Press.

McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods*, *37*(4), 547–559.

Mestre, D. R. (2015). On the usefulness of the concept of presence in virtual reality applications. *The Engineering Reality of Virtual Reality 2015*, *9392*, 93920J.

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

Meteyard, L., Cuadrado, S. R., Bahrami, B., & Vigliocco, G. (2012). Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *48*(7), 788–804.

Meyer, K., & Damasio, A. (2009). Convergence and divergence in a neural architecture for recognition and memory. *Trends in Neurosciences*, *32*(7), 376–382.

Mollo, G., Pulvermüller, F., & Hauk, O. (2016). Movement priming of EEG/MEG brain responses for action-words characterizes the link between language and action. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *74*, 262–276.

Moore, K., Wiederhold, B. K., Wiederhold, M. D., & Riva, G. (2002). Panic and agoraphobia in a virtual world. *CyberPsychology & Behavior*, *5*(3), 197–202.

Morales, L., Paolieri, D., Dussias, P. E., Kroff, J. R. V., Gerfen, C., & Bajo, M. T. (2016). The gender congruency effect during bilingual spoken-word recognition. *Bilingualism*, *19*(2), 294–310.

Moreno, I., Vega, M. D., León, I., Bastiaansen, M. C. M., Lewis, A. G., & Magyari, L. (2015). Brain dynamics in the comprehension of action-related language: A time frequency analysis of mu rhythms. *NeuroImage*, *109*(April), 50–62.

Morgan-Short, K., & Tanner, D. (2013). Event-related potentials (ERPs). *Research Methods in Second Language Psycholinguistics*, 127–152.

Moseley, R. L., Mohr, B., Lombardo, M. V., Baron-Cohen, S., Hauk, O., & Pulvermüller, F. (2013). Brain and behavioral correlates of action semantic deficits in autism. *Frontiers in Human Neuroscience*, *7*.

Newman, A. J., Tremblay, A., Nichols, E. S., Neville, H. J., & Ullman, M. T. (2012). The Influence of Language Proficiency on Lexical Semantic Processing in Native and Late Learners of English. *Journal of Cognitive Neuroscience*, *24*(5), 1205–1223.

Niccolai, V., Klepp, A., Weissler, H., Hoogenboom, N., Schnitzler, A., & Biermann-Ruben, K. (2014). Grasping hand verbs: Oscillatory beta and alpha correlates of action-word processing. *PloS One*, *9*(9), e108059.

Ojima, S., Nakata, H., & Kakigi, R. (2005). An ERP Study of Second Language Learning after Childhood: Effects of Proficiency. *Journal of Cognitive Neuroscience*, *17*(8), 1212–1228.

O'Neil, Lagarrigue, Newman & Frenck-Mestre (2016). Unpublished manuscript.

Öttl, B., Dudschig, C., & Kaup, B. (2017). Forming associations between language and sensorimotor traces during novel word learning. *Language and Cognition*, *9*(1), 156–171.

Papeo, L., Vallesi, A., Isaja, A., & Rumiati, R. I. (2009). Effects of TMS on different stages of motor and non-motor verb processing in the primary motor cortex. *PloS One*, *4*(2), e4508.

Papesh, M. H. (2015). Just out of reach: On the reliability of the action-sentence compatibility effect. *Journal of Experimental Psychology. General*, *144*(6), e116-141.

Patterson, K., & Lambon Ralph, M. A. (2016). Chapter 61—The Hub-and-Spoke Hypothesis of Semantic Memory. In G. Hickok & S. L. Small (Eds.), *Neurobiology of Language* (pp. 765–775). Academic Press.

Peeters, D. (2019). Virtual reality: A game-changing method for the language sciences. *Psychonomic Bulletin & Review*, *26*(3), 894–900.

Peeters, D., & Dijkstra, T. (2018). Sustained inhibition of the native language in bilingual language production: A virtual reality approach. *Bilingualism: Language and Cognition*, *21*(5), 1035–1061.

Peterson, M. (2010). Computerized games and simulations in computer-assisted language learning: A meta-analysis of research. *Simulation & Gaming*, *41*(1), 72–93.

Pfurtscheller, G., & Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: Basic principles. *Clinical Neurophysiology*, *110*(11), 1842–1857.

Pineda, J. A. (2005). The functional significance of mu rhythms: Translating "seeing" and "hearing" into "doing." *Brain Research. Brain Research Reviews*, *50*(1), 57–68.

Price, A. R., Peelle, J. E., Bonner, M. F., Grossman, M., & Hamilton, R. H. (2016). Causal Evidence for a Mechanism of Semantic Integration in the Angular Gyrus as Revealed by High-Definition Transcranial Direct Current Stimulation. *The Journal of Neuroscience*, *36*(13), 3829–3838.

Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences*, *22*(2), 253–336.

Pulvermüller, F. (2002). A brain perspective on language mechanisms: From discrete neuronal ensembles to serial order. *Progress in Neurobiology*, *67*(2), 85–111.

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews. Neuroscience*, *6*(7), 576–582.

Pulvermüller, F. (2010). Brain embodiment of syntax and grammar: Discrete combinatorial mechanisms spelt out in neuronal circuits. *Brain and Language*, *112*(3), 167–179.

Pulvermüller, F. (2013). How neurons make meaning: Brain mechanisms for embodied and abstract-symbolic semantics. *Trends in Cognitive Sciences*, *17*(9), 458–470.

Pulvermüller, F. (2018). Neurobiological Mechanisms for Semantic Feature Extraction and Conceptual Flexibility. *Topics in Cognitive Science*, *10*(3), 590–620.

Ranalli, J. (2008). Learning English with The Sims: Exploiting authentic computer simulation games for L2 learning. *Computer Assisted Language Learning*, *21*(5), 441–455.

Raposo, A., Moss, H. E., Stamatakis, E. A., & Tyler, L. K. (2009). Modulation of motor and premotor cortices by actions, action words and action sentences. *Neuropsychologia*, *47*(2), 388–396.

Repetto, C. (2014). The use of virtual reality for language investigation and learning. *Frontiers in Psychology*, *5*.

Repetto, C., Cipresso, P., & Riva, G. (2015). Virtual action and real action have different impacts on comprehension of concrete verbs. *Frontiers in Psychology*, *6*.

Rodríguez-Fornells, A., Cunillera, T., Mestres-Missé, A., & de Diego-Balaguer, R. (2009). Neurophysiological mechanisms involved in language learning in adults. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1536), 3711–3735.

Rommers, J., Dickson, D. S., Norton, J. J. S., Wlotko, E. W., & Federmeier, K. D. (2017). Alpha and theta band dynamics related to sentential constraint and word expectancy. *Language, Cognition and Neuroscience*, *32*(5), 576–589.

Sakreida, K., Scorolli, C., Menz, M. M., Heim, S., Borghi, A. M., & Binkofski, F. (2013). Are abstract action words embodied? An fMRI investigation at the interface between language and motor cognition. *Frontiers in Human Neuroscience*, *7*.

Sá-Leite, A. R., Fraga, I., & Comesaña, M. (2019). Grammatical gender processing in bilinguals: An analytic review. *Psychonomic Bulletin & Review*, *26*(4), 1148–1173.

Sato, M., Mengarelli, M., Riggio, L., Gallese, V., & Buccino, G. (2008). Task related modulation of the motor system during language processing. *Brain and Language*, *105*(2), 83–90.

Saussure (1857-1913), F. de. (1916). *Cours de linguistique générale*.

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, *3*(3), 417–424.

Shapiro, L. (2011). *Embodied cognition* (pp. xiii, 237). Routledge/Taylor & Francis Group.

Shebani, Z., & Pulvermüller, F. (2018). Flexibility in Language Action Interaction: The Influence of Movement Type. *Frontiers in Human Neuroscience*, *12*.

Simmons, W. K., & Barsalou, L. W. (2003). The similarity-in-topography principle: Reconciling theories of conceptual deficits. *Cognitive Neuropsychology*, *20*(3), 451–486.

Simmons, W. K., & Martin, A. (2009). The anterior temporal lobes and the functional architecture of semantic memory. *Journal of the International Neuropsychological Society*, *15*(5), 645–649.

Skulmowski, A., & Rey, G. D. (2018). Embodied learning: Introducing a taxonomy based on bodily engagement and task integration. *Cognitive Research: Principles and Implications*, *3*.

Sundqvist, P. (2009). *Extramural English Matters: Out-of-School English and Its Impact on Swedish Ninth Graders' Oral Proficiency and Vocabulary*.

Sylvén, L. K., & Sundqvist, P. (2012). Gaming as extramural English L2 learning and L2 proficiency among young learners*. *ReCALL*, *24*(3), 302–321.

Tellier, M. (2008). The effect of gestures on second language memorisation by young children. *Gesture*, *8*(2), 219–235.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action* (pp. xxiii, 376). The MIT Press.

Thompson-Schill, S. L. (2003). Neuroimaging studies of semantic memory: Inferring “how” from “where.” *Neuropsychologia*, *41*(3), 280–292.

Tirado, C., Khatin-Zadeh, O., Gastelum, M., Leigh-Jones, N., Marmolejo-Ramos, F. (2018). The strength of weak embodiment. *International Journal of Psychological Research*, *11*(2), 77–85.

Tomasello, R., Garagnani, M., Wennekers, T., & Pulvermüller, F. (2018). A Neurobiologically Constrained Cortex Model of Semantic Grounding With Spiking Neurons and Brain-Like Connectivity. *Frontiers in Computational Neuroscience*, *12*.

Török, A., Sulykos, I., Kecskés-Kovács, K., Persa, G., Galambos, P., Kóbor, A., Czigler, I., Csépe, V., Baranyi, P., & Honbolygó, F. (2014). Comparison between wireless and wired EEG recordings in a virtual reality lab: Case report. *2014 5th IEEE Conference on Cognitive Infocommunications (CogInfoCom)*.

Tromp, J., Peeters, D., Meyer, A. S., & Hagoort, P. (2018). The combined use of virtual reality and EEG to study language processing in naturalistic environments. *Behavior Research Methods*, *50*(2), 862–869.

Tschentscher, N. (2017). Embodied Semantics: Embodied Cognition in Neuroscience. *German Life and Letters*, *70*(4), 423–429.

van Elk, M., van Schie, H. T., Zwaan, R. A., & Bekkering, H. (2010). The functional role of motor activation in language processing: Motor cortical oscillations support lexical-semantic retrieval. *NeuroImage*, *50*(2), 665–677.

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

Wakefield, E. M., & James, K. H. (2011). Effects of sensori-motor learning on melody processing across development. *Cognition, Brain, Behavior : An Interdisciplinary Journal*, *15*(4), 505–534.

Wamain, Y., Sahaï, A., Decroix, J., Coello, Y., & Kalénine, S. (2018). Conflict between gesture representations extinguishes μ rhythm desynchronization during manipulable object perception: An EEG study. *Biological Psychology*, *132*, 202–211.

Warrington, E. K., & McCarthy, R. (1983). Category specific access dysphasia. *Brain*, *106*(4), 859–878.

Weber-Fox, C., & Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience*, *8*(3), 231–256.

Willems, R. M., Toni, I., Hagoort, P., & Casasanto, D. (2009). Neural Dissociations between Action Verb Understanding and Motor Imagery. *Journal of Cognitive Neuroscience*, *22*(10), 2387–2400.

Wilson, R. A., & Foglia, L. (2017). Embodied Cognition. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2017). Metaphysics Research Lab, Stanford University.

Young, M. F., Slota, S., Cutter, A. B., Jalette, G., Mullin, G., Lai, B., Simeoni, Z., Tran, M., & Yukhymenko, M. (2012). Our Princess Is in Another Castle: A Review of Trends in Serious Gaming for Education. *Review of Educational Research*, *82*(1), 61–89.

Zwaan, R. A., and Madden, C. J., D. (2005). *“Embodied sentence comprehension,” in Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thinking*. Cambridge University Press.

Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: Motor resonance in language comprehension. *Journal of Experimental Psychology: General*, *135*(1), 1–11.