UNIVERSITÀ CATTOLICA DEL SACRO CUORE

Sede di Milano

Dottorato di ricerca in Psicologia

Ciclo XXXV

S.S.D. M-PSI/01

Space in Language: Embodied Evidence of Space-Language Interaction

Coordinatore:

Ch.mo Prof. Camillo Regalia

Tesi di Dottorato di:

Cosimo Tuena

N. Matricola: 4912271

Anno Accademico 2021/2022

Summary

Chapter 1: Language and Space from an Embodied Perspective

The classic cognitive science perspective has typically regarded the mind as an abstract information processor with little theoretical relevance to its connections to the outside world (M. Wilson, 2002). In the 70s and 80s, such a view was endorsed by modularism and the physical symbol system hypothesis (Fincher-Kiefer, 2019). The conceptualization of the mind as a modular system saw the cognitive processes as separated and with little interaction among each other; the sensorimotor system was conceptualized with an input/output function and little role in 'central' cognition (Fodor, 1983). The physical symbol system hypothesis stated that knowledge was amodal and meaning emerged from the manipulation of a symbol and the relations to other symbols (e.g., other words). According to standard cognitive theories (Fincher-Kiefer, 2019), knowledge is stored separately from the brain's modal systems for perception (e.g., vision, audition), action (e.g., movement, proprioception), and introspection (e.g., mental states, affect). That is to say that cognitive representations are converted by the mind into modality-independent (i.e., abstract) representations of any experiences, internal states, or knowledge. Nevertheless, some psychologists, such as William James, Jean Piaget, James Gibson, George Lakoff, and Mark Johnson, pointed out the importance of the modality-dependent representations and the sensorimotor system for 'central' cognition (Gibson, 1979; James et al., 1890; Lakoff & Johnson, 1980; Piaget, 1953). For instance, Piaget put as the first step of cognitive development the sensorimotor exploration of the child with the world (Piaget, 1953); similarly, Gibson (Gibson, 1979) showed that perception is strictly related to the action possibilities (affordances) that we perceive in the environment that surrounds us. Again, Lakoff and Johnson (Lakoff & Johnson, 1980) proposed that abstract concepts and language metaphors are rooted in the body and the sensorimotor experience that we have with the external world. Lastly, the roots of the standard cognitive theories of the mind were shaken by the so-called 'symbol-grounding' problem (Fincher-Kiefer, 2019). The problem proposed that symbols (i.e., words) that have no external connections can have no meaning, and finding words' meaning by their relations to other symbols leads to an endless search (Glenberg, 2015). The only way to break the loop is to ground symbols in the bodily experience.

In the past twenty years, this view of the mind has taken the name of embodied cognition (EC). This view has attained high visibility since there is growing theoretical and experimental evidence that cognition can be better understood in relation to the body and environment. Fincher-Kiefer (Fincher-Kiefer, 2019) proposes three main approaches to EC. Glenberg's view (Glenberg, 2015) assigns to the body a central role in knowledge. Barsalou's view (Barsalou, 1999) where representation is supported by bodily experiences and the sensorimotor simulation. Lakoff and Johnson's view (Lakoff & Johnson, 1980) proposes that language supports our representation of knowledge.

Despite a high overlap and consensus on the key role of the body and environment for cognitive representations, within the EC scientific community differences and criticisms have been raised. First, EC has led to different views on how representations are supported by 'central' and 'peripheral' systems (Barsalou, 1999, 2008; Glenberg, 2010; M. Wilson, 2002). The main division is concerning the existence of an internal representation or not. According to the non-representational embodied view, the sensorimotor system provides the necessary feedback also from the action-environment interaction (Chemero, 2011; Gibson, 1979). Researchers that endorse this view hint that a cognitive representation is not necessary to support particular interactions with the environment guided by perception (A. D. Wilson & Golonka, 2013). Conversely, the representational embodied view accepts the existence of a representation that is multimodal and sustained by sensorimotor and abstract information (Barsalou, 1999). Such a representation grounded in the sensorimotor system is crucial to support different psychological functions, such as memory, language, or knowledge (Barsalou, 2008; Dijkstra & Post, 2015; Gallese & Lakoff, 2005; Körner et al., 2016). Second, Wilson (M. Wilson, 2002) provided a critical overview of some common assumptions within the EC framework.

A cornerstone of EC is that cognition is situated and occurs during task-relevant inputs and outputs. However, research has shown that the embodied mind can work also offline without direct interaction with the environment. In addition, cognition occurs in real-time (time-pressured) and this implies creating cheap and efficient cognitive models; however, humans may fail to operate under time constraints and, when the possibility is given, offline representation/cognition is preferred. On certain occasions, the 'cheap and efficient' time-pressured representation is the only one available. In this case, two ways can be used to help us: prior learning or altering the environment to reduce cognitive effort. A radical view here suggests that cognition is not restricted to the mind but is distributed and includes the body and the environment. Although this view has been endorsed by some theorists, the fact that the causes of behavior are distributed is not sufficient to study a distributed system. Rather, it is important to explore the organization and function of such a complex interaction.

Within the EC framework, some researchers have tried to explain cognitive phenomena, like memory and visual perception, as aiming to action (cognition is for action). However, this is not always true. Instead, the question is how cognition sustains action and how information stored in an unspecific way can be used for it. Lastly, it is crucial to understand that in parallel to online (or situated) aspects of EC, the systems for perception and action can operate offline to run simulations without the actual use of the body (Barsalou, 2008). Evidence of this comes from EC studies on memory, reasoning, language, and abstract concepts, which show that sensory and/or motor information can be used offline for different cognitive activities (offline cognition is body-based). Another criticism of the view that sees the body as necessary for representation is that cognition often occurs independently of the body. The term 'grounded cognition' (Barsalou, 2008) refers to the idea that cognition is typically grounded in a variety of ways, such as simulations, situated action, and, on occasion, bodily states.

In addition, the role of the body in information processing (i.e., embodiment) has demonstrated that cognition can be grounded in several manners. Embodiment can be defined as a phenomenon in which the body, its sensory states, morphology, or mental representations affect information processing (Körner et al., 2016). According to the review by Körner and colleagues (Körner et al., 2016) the body and the sensorimotor system can have three distinct effects on cognition. First, they can affect a person's mood, sentiments, or information processing directly (direct state induction). For instance, injecting Botulinum Toxin A inhibits the corrugator muscle and is correlated with positive affect in accordance with the facial feedback hypothesis (Lewis & Bowler, 2008). Second, they can influence the mental contents rather by altering the ease with which specific information comes to mind (modal priming). For instance, holding a heavy or light object affects the subsequent importance estimation of an object or topic (Jostmann et al., 2009; Schneider et al., 2011). Findings in the so-called modal priming are also supported by the strong link between bodily and sensorimotor experiences and abstract concepts using metaphors (Gallese & Lakoff, 2005; Lakoff & Johnson, 1980). Third, they may cause compatibility issues with concurrent automatic simulations (sensorimotor simulation). This can be defined as the re-enactment of perceptual, motor, and introspective states acquired during a given experience with the world, body, and mind (Barsalou, 2008). Later, when the object (e.g., an event, a chair) of the experience is re-experienced, a multimodal representation of it arises as it was captured. Such simulation can be triggered by object affordances or by words that imply sensorimotor states (Glenberg, 2010; Körner et al., 2016). In this work and the following sections, I will focus on the latter view of EC, which supports the presence of a multimodal representation to support cognition. In particular, evidence in favour of the embodied nature of two cognitive domains (i.e., language and spatial cognition) will be examined.

Embodied Language

Classic cognitive linguistic theories supported the notion that language and its components (e.g., phonology, phonetics, semantic meaning) were non-perceptual. Language conveys meaning by combining syntactic rules with abstract, amodal, and arbitrary symbols (i.e., words) (Chomsky, 1980; Fodor, 2000). The early proposal that language is embodied and situated came from studies on abstract concepts and metaphors and the link between language components (e.g., syntax, semantics) and components of experience (e.g., spatial relations) (Lakoff & Johnson, 1980; Talmy, 1983). In the 90s, thanks to the discovery of the mirror neuron system (Rizzolatti & Craighero, 2004) in the premotor and motor cortex, further studies enabled to support the notion that human language comprehension is rooted in the sensorimotor system (Tettamanti et al., 2005). Despite body and language comprehension seem to be apart, findings support that these two domains are inherently linked. This statement comes from different EC approaches to language comprehension (Barsalou, 2008).

First, during language comprehension humans create a model of the situation (situated models) (Zwaan & Radvansky, 1998). Such models can have spatial properties (e.g., point of view of action) and multimodal representations, such that specific spatial perspective and/or motor simulation can arise from situated models of, for instance, a sentence involving multiple characters (Beveridge & Pickering, 2013). Second, according to the classic theories of language comprehension, the input of a sentence is converted into a prepositional representation; however, Zwaan and colleagues (Zwaan et al., 2002) showed that the representation of the content of a sentence is composed of perceptual symbols (Barsalou, 1999) that match the implicit (perceptual) simulation conveyed by the sentence. The perceptual symbol system proposed by Barsalou (1999) supports the formation of multimodal representation of the experience. Knowledge is represented through modal perceptual symbols that sustain cognition, which according to Barsalou is inherently perceptual. After an experience with an object (e.g., a chair), an analogue of modal symbol is stored in memory and used to extract all the relevant sensorimotor or perceptual elements that were coded when experiencing with a chair.

Then, according to the indexical hypothesis (Glenberg & Gallese, 2012), action language is understood by creating a motor simulation of the action represented in the sentence (e.g., meaning arises from action). Several pieces of evidence endorse that the motor information and brain areas are crucial to simulate and understand action language (Birba et al., 2017; Glenberg & Kaschak, 2002; Papeo et al., 2011; Repetto et al., 2013). Lastly, when the content of an experience or sentence is linked to emotional states, the subject simulates such states (affective simulation). For instance, the use of the first-person pronoun enhances the immersion during narrative reading (Hartung et al., 2016).

Nevertheless, despite a common agreement among EC researchers, there are differences between the two complementary (and not contradictory) views on language comprehension between Glenberg and Barsalou's view (Fincher-Kiefer, 2019). The former affirms that comprehension arises from the simulation of past experiences, imagination, and in general the perceptual symbols. The Glenberg and colleagues' approach (Glenberg et al., 2008; Glenberg & Gallese, 2012; Glenberg & Kaschak, 2002) states that comprehension is grounded in systems for perception and action planning and meaning is conveyed and supported by the neural systems that are used to plan and execute actions. In their seminal paper, Glenberg and Kaschak (Glenberg & Kaschak, 2002) reported the so-called action-sentence compatibility effect (ACE). In their experiment, they asked participants to read concrete ("*Courtney handed you the notebook/You handed Courtney the notebook*", direction of the action toward the body/away from the body respectively), abstract ("*Liz told you the story/You told Liz the story*", toward/away), imperative ("*Open the drawer/Close the drawer*", toward/away) and non-sense sentences while pressing a central button on a horizontal keyboard. Two response conditions were provided, if the sentence made sense, they had to press on the keyboard a button far from the body if the sentence did not make sense they pressed a button near to the body, or vice versa. They found an interaction (button distance by direction of the sentence) for imperative, concrete, and abstract transfer sentences, where reading times were shorter for away sentences when participants had to plan the button response far from the body compared to the near button and for toward sentences when participants had to plan to press the button near the body in contrast to the far button. In the attempt to study the involvement of the motor neural pathways during action sentence comprehension, Glenberg and co-authors (Glenberg et al., 2008) used single-pulse transcranial magnetic stimulation (TMS) and motor-evoked potentials (MEP) while reading concrete and abstract transfer vs. no transfer sentences and non-sense sentences. They found that MEP were larger during the reading in proximity of the verb of the sentence compared to the end, suggesting the recruitment of the motor system rather than a process of imagery, and larger MEP for abstract and concrete transfer compared to no transfer sentences. This suggests that an ‘action schema’ (motor simulation) is used during action language comprehension.

Despite the drastic impact of such findings, recent studies failed to replicate the interaction (button distance by direction of the sentence) that sustain ACE. Intriguingly, an ACE replication study (Díez-Álamo et al., 2020) it was found that toward sentences were processed faster and remembered better compared to away sentences. The authors motivated such a result by highlighting the perceptual and attentional relevance of objects that are approaching towards the body. They called this the linguistic looming effect (LLE), where linguistic information, as the physical one, is sensitive to looming objects. In a large multicenter replication project (Morey et al., 2022) different research teams administered the ACE paradigm to English and non-English participants. They failed to replicate the ACE and found instead the same main effect described above.

Another question is to what extent and what sensorimotor areas are recruited during comprehension of sentences and words. Tettamanti and co-authors (Tettamanti et al., 2005) found that listening to action-related sentences of different body parts (e.g., leg “*I kick the ball*”; mouth “*I bite the apple*”; hand “*I grasp the knife*”) activates a left frontoparietal network that includes the motor brain regions specific to the body parts involved in the sentences. In the Masson and colleagues’ (Masson et al., 2008) experiment participants were trained to make a hand action in response to a visual cue while listening to a sentence to test the evocation of motor representations during sentence comprehension. The results showed that the motor activation that underlies a sentence (e.g., "*the lawyer kicked aside the calculator*") is not specific, instead, it includes a sensorimotor representation associated with the object (e.g., a calculator) in our past experiences. Thus, it is possible that the underlying simulation is not restricted to the motor system (Tettamanti et al., 2005) but activates other perceptual representations associated with the action described (Barsalou, 1999). Another interesting question is if the motor resonance occurs only within the central nervous system or if it spreads in the peripheral one. Repetto and co-authors addressed this question in their study (Repetto et al., 2021). Participants learned written words, an image illustrating the word's meaning, performed a gesture representing the word's meaning, or observed and performed the gesture. Free recall and visual word recognition tasks were used to assess word retention. The electromyogram of the forearm muscle was then recorded as the participants completed a visual word recognition task. Recognition results showed that words learned through self-performed gestures elicited more muscle activation than words encoded in other training conditions. The authors concluded that memory recognition is supported also by peripheral re-activation in addition to the central motor simulation.

The study of the motor system contribution is also supported by studies on neurological conditions. For instance, De Scalzi and co-authors (De Scalzi et al., 2015) showed that sentence comprehension and motor simulation of the action are preserved in older people and individuals that suffer from Alzheimer's disease (AD). The authors found a preserved ACE in these two populations suggesting that in AD the preserved motor abilities enable language simulation and comprehension. Conversely, influent research on Parkinson's disease (PD) showed that in neurological conditions where the motor system (basal ganglia and frontostriatal system) is affected by neurodegeneration, the simulation of action language is hampered (for a review, see Birba et al., 2017). For instance, Cardona and colleagues (Cardona et al., 2014) showed that individuals affected by PD have a disrupted ACE, whereas healthy controls and patients affected by myelitis did not. The results endorse, for the ACE paradigm, the representational embodied view and discard a non-representational embodied view.

Barsalou and colleagues' view (Barsalou, 1999; Zwaan et al., 2002) demonstrates that language comprehension is supported by re-enactments of a sensorimotor experience. For instance in Zwaan and co-authors' study (Zwaan et al., 2002), participants had to decide if an item represented in a picture was present in a previously presented sentence. The authors predicted that when the implied orientation of the object in the sentence matched the object's orientation in the picture, responses would be faster than when the implied and pictured orientations differed. For instance, response times were slower when "*the ranger saw the eagle in the sky*" is followed by a picture of an eagle with folded wings compared to a picture of an eagle with outstretched wings. Results support the authors' predictions and the importance of perceptual simulation during sentence comprehension. Willems and colleagues (Willems et al., 2010) used functional magnetic resonance imaging to study motor cortex activation in right- and left-handers of manual (e.g., to grasp) vs- non-manual action verbs. During the lexical decision task (i.e., if the word exists or not), premotor areas corresponding to the contralateral handedness side were activated, whereas during the imagery task the activation included the motor cortex. This study hints that semantic decision and imagery have distinct brain regions that are used to comprehend according to body characteristics. Using only right-handed participants, Repetto and colleagues (Repetto et al., 2013) found that repeated TMS over the hand motor areas inhibits comprehension. In their study, the authors presented words describing a hand-related action (e.g., to grasp) or an abstract (e.g., to imagine) verbs. RTs to judge whether the word was concrete or abstract were slower when the rTMS was applied to the left-hand motor cortex of the right-handed participants compared to the right-hand motor cortex stimulation and this was true only for concrete verbs.

Most of the EC research has focused, within the domain of language, on semantic meaning (Fincher-Kiefer, 2019). However, some authors suggest that also language components related to syntax and phonology can be seen as embodied. More specifically, Broca's area, which is responsible for the production of language, is linked also to syntax and phonology/phonetics (Fogassi & Ferrari, 2008). For instance, listening to words activates the corresponding premotor cortex that sustain the production of the phoneme and the Broca's area (Rizzolatti & Craighero, 2004). Indeed, there is mounting evidence that phonology is formed, stored, and transmitted through embodied sound production learning, moulded not only by environmental inputs but also by the experience of making and identifying sounds (Nathan, 2017). However, if on the one hand phonetics (sound perception/production) is thought to be embodied (i.e., is based on a sensorimotor simulation), on the other hand, EC scientists are still debating whether this could be for phonology. Indeed, strong positions on embodiment assert that both phonetics and phonology are embodied, whereas other accounts suggest that phonology is embodied or abstract depending on the aspect to be considered (Berent & Platt, 2022). According to Berent and Platt (2022), motor simulation occurs at different stages. In particular, motor simulation occurs during sound speech perception (phonetics) but also during phonological lexical associations, conversely, the phonological structure (e.g., syllable) relies primarily on algebraic rules (cf.; Chomsky, 1980).

Despite some researchers endorses this view, others suggest that also syntaxis is partially supported by embodiment and sensorimotor simulation. Different studies from psycholinguistics state that prosody is crucial for acquisition, production, and comprehension of language (Kreiner & Eviatar, 2014). Consequently, syntax is grounded via the simulation of prosodic patterns (for a review, see Kreiner & Eviatar, 2014). Given that the brain has functional links between syntax and prosody, it may be easier to anchor syntactic processes in the brain via prosody's structure than without it. According to the oscillation-based model (Giraud & Poeppel, 2012) neural oscillations in the low frequency delta band, which have been linked to prosodic perception and production, may be evoked to simulate prosody in the absence of external prosodic input. This type of simulation can be used as a foundation for the implementation of syntax.

This set of findings strongly suggests that language comprehension at the semantic level is embodied, and this occurs through different EC mechanisms, conversely, at the phonological/phonetics level, only a part of language is embodied.

Embodied Space

The classic cognitive theory (Huffman & Ekstrom, 2019; O'Keefe & Nadel, 1978; Wolbers et al., 2011) claims that space representations are amodal. This representation is also called an allocentric representation of the space (Burgess, 2008; O'Keefe & Dostrovsky, 1971). An allocentric map is independent of our body location and point of view and is abstract. Such classic modality-independent theory proposes that this cognitive map of the environment is fundamentally separated from the information in which was encoded. Indeed, humans learn about the space through both cognitive (e.g., spatial reasoning, route decision, previous memories) and bodily (e.g., motor commands, vision, vestibular information, proprioception) self-generated/active information (Chrastil & Warren, 2012).

However, in addition to an allocentric representation of the space, it is possible to have an egocentric map of the environment (Burgess, 2008). This representation is dependent on the body, our location and point of view, and it is based on sensorimotor information. Hence, spatial information can be coded with two spatial frames of reference: the allocentric and egocentric representations. Both can be built by using bodily and environmental (e.g., landmark, boundaries) information (Barry & Burgess, 2014; Lester et al., 2017).

The model proposed by Byrne and colleagues (Byrne et al., 2009) elucidates how these two representations interact and how spatial memory is formed and retrieved. We acquire information from the space from an egocentric point of view sustained by the parietal cortex, then this is translated, through the medial parietal region, and stored in long-term memory into an allocentric map in the hippocampus and medial temporal lobe. When retrieved, the allocentric information is back-transformed into a body-based representation.

Despite great consensus on the existence of an abstract cognitive map of the environment, researchers are debating whether the sensorimotor system impacts the encoding and retrieval of this map (Huffman & Ekstrom, 2021; Steel et al., 2021). Following this statement, the modality-dependent hypothesis claims that representations are not necessarily abstract and can be influenced by the mode of acquisition (Steel et al., 2021).

Thanks to the introduction of virtual reality (VR), alone or combined with neurophysiological measures, spatial cognition and navigation research have provided important advances in understanding the cognitive and neural mechanisms of these domains in humans (Bohil et al., 2011; Burgess et al., 2001). Thanks to its multisensory and bodily potential, VR is a feasible tool to understand the role of the body in spatial cognitive maps of the environment (Tuena, Serino, et al., 2021).

On one side, Huffman and Ekstrom (2019) studied how different degrees of bodily information influence spatial memory (pointing task). In their task participants navigated a large-scale environment with a treadmill and visor, with joypad and visor or with desktop PC and joypad. There was not any effect of the navigation condition nor at the behavioural neither at the neurophysiological levels. They concluded that visual information may play a crucial role in large-scale environments compared to other body-related information (e.g., vestibular information of the treadmill and visor group). Again, Tuena and colleagues (Tuena et al., 2017) in a pilot study tried to modify body engagement during an ‘active’ vs. ‘passive’ navigation task in a virtual city for the assessment of episodic memory. They manipulated bodily information through a full-embodiment condition (participants wore a 3D visor and a Kinect camera tracked in-place leg movement to enable a virtual walk), a medium-embodiment condition (they simulated an in-place walk while pre-registered navigation was shown in the 3D visor), and a low-embodiment condition (they passively watched the pre-recorded video with the visor). They found no effect of such conditions on subsequent episodic and spatial memory verbal recall and visual recognition. However, the authors found that the sense of presence (i.e., the illusion of being located in the virtual world; Riva, Castelnuovo, and Mantovani 2006) was higher in the first condition compared to the other two.

On the other side, a systematic review (Tuena et al., 2019) of 31 studies showed that active navigation (i.e., use of motor information in addition to vision) compared to passive navigation (i.e., use of visual information only) in VR enhances subsequent memory recall of spatial information under the egocentric and allocentric representation. This could indicate that the active use of motor information during virtual navigation improves memory encoding and enriches the memory traces with additional information at retrieval, namely the *virtual enactment effect*. This mechanism is similar to the *enactment effect* (Engelkamp, 1998; Engelkamp et al., 2004), which is described as the enhancement given by encoding sentences (e.g., "*open the bottle*") with the execution of actions that depict the information conveyed by the phrase, instead of only reading the sentence or watching someone else acting. Interestingly the study by Plancher and co-authors (Plancher et al., 2012) showed that active navigation compared to the passive exploration enhanced allocentric spatial memory in healthy, amnestic mild cognitive impairment, and AD individuals. They suggested that the motor information used during the active condition provided additional memory traces that helped the patients with memory problems to retrieve hippocampal-based information. In addition to the role of the sensorimotor system, other researchers studied the impact of affordances during spatial navigation. Researchers have found that environmental information (i.e., discrete landmarks, boundaries) can be processed as affordances that facilitate or limit navigation (Julian et al., 2018). This is further enhanced by affective states associated with a given landmark. For what concerns the role of the affective states, in the study of Ruotolo and co-authors (Ruotolo et al., 2018) participants passively watched a route with positive, negative, or neutral images (landmarks) taken from the International Affective Picture System (IAPS; Lang, 1995) at turning points. The authors found that participants who watched the path with positive landmarks were better at ordering the images and at drawing a map of the path. Conversely, participants in the negative landmark condition rated the route as longer than the positive and neutral ones and took more time to mentally travel between landmarks. In a similar study by Piccardi and co-authors (Piccardi et al., 2020), participants were asked to perform a real-world navigation task (learning and recalling a path in the Walking Corsi Test) or to do a non-embodied task ('paper and pencil' path drawing and landmark recognition) to study the effect of emotional (positive/negative, high arousal/low arousal) and neutral landmarks (i.e., photos from the IAPS; Lang 1995). Results showed that the affective landmarks improved path learning in the embodied condition, but the recall performance was the same in the two conditions.

In terms of actions, Morganti (Morganti, 2018) showed that the VR version of a traditional test (Money Road Map test) was more effective than the 'paper and pencil' version in providing egocentric inputs useful for orientation. In this test, the participant is required to describe turns (i.e., left-right) while watching a path on a city-like map and cannot turn the map around to match her/his perspective with the one on the map. Conversely, participants in the VR version of the test did not have to re-locate continuously on the map, instead, landmarks-based turns were done on the body axis without changing the perspective (i.e., by moving the point of view with the joypad). According to the authors, VR gives an enactive spatial representation (sensorimotor coupling with agent's actions in correspondence of landmark) that is different from typical cognitive tests. Indeed, VR enables individuals to use landmarks as affordances to plan navigation. While the egocentric frame is naturally processed within an enacted and embodied approach (it is action-oriented) (Borghi & Barsalou, 2021), research also demonstrates that the allocentric landmarks representation could be embodied (König et al., 2019). In their study, König and colleagues (2019) evaluated the allocentric map of participants' hometown by assessing unitary coding (angular difference between the orientation of a well-known building or street and true north) and binary coding (angular difference between the orientation of two well-known houses or two well-known streets; pointing from one well-known building to another well-known building) under spontaneous (time-pressured, 3s to respond) and cognitive reasoning (no time-limit) conditions. The prediction is that building (but not streets) binary coding is accessed automatically to generate navigation behavior and action, whereas unitary coding requires cognitive processing. The authors showed that, when time-pressured, the retrieval of buildings relations yielded better performances compared to building unitary coding. Conversely, when spontaneous action-related information is suppressed by cognitive reasoning, an inverse pattern emerged both for buildings and streets' cardinal orientation.

The authors concluded that allocentric information about the orientation and position of landmarks can likewise be coded within an enacted and embodied representation. The importance of action-landmark coupling is highlighted also by studies on aging. Cogné and co-authors (Cogné et al., 2018) administered a VR navigation task and found that landmarks (e.g., elements in the environment associated with navigation decision-making and motor commands – "*at the church, I need to go right*") and directional cues improved mild cognitive impairment and AD spatial recall. The presence of such visual cues helps the construction of a spatial map of the egocentric type (sequence of landmarks and route between landmarks).

There is initial evidence that the cognitive map of the space can be affected by action, perception, and internal states. However further studies need to address the limits and conflicting results reported in the literature. Different hypotheses have been made regarding the impact of VR interface (non-immersive, semi-immersive, or immersive), the role of the spatial frame of reference (egocentric vs. allocentric), and environment scale (large vs. peripersonal) (Huffman & Ekstrom, 2021; Steel et al., 2021). Taube and colleagues (Taube et al., 2013) argued that current neurophysiological and behavioral results coming from the non-immersive interfaces that use only visual or visual combined with motor information (e.g., joypad) give us incomplete results as such condition is not comparable to the more complex role of the body during real-world navigation. Immersive VR setup or real-world task would better explain the role of bodily information during the encoding and retrieval of spatial cognitive maps. Then, the impact of the body on spatial representations might also depend on the type of spatial frames of references implied in the task, where egocentric representation is mostly affected by sensorimotor information whereas the allocentric one is abstract and modality-independent. Lastly, the scale (e.g., a big city vs. near-space representation) of the environment could be influenced in different ways by the body, with the latter relying more than the former on the body.

Space and Language in Interaction

As described in the first section, modularism (Fodor, 1983) asserted that the human mind is, at least partially, divided into modules that operate more or less independently of each other. Such a view has been challenged by EC and more recent evidence that shows how cognitive domains are intertwined and sustained by common brain networks (Fincher-Kiefer, 2019; Fogassi & Ferrari, 2008; Spreng et al., 2008; Tettamanti et al., 2005; Vukovic & Shtyrov, 2017). Hence, multiple cognitive domains and brain systems support the formation and retrieval of multimodal and complex cognitive representations (Barsalou, 1999, 2008). About the topic of this thesis, recent studies demonstrate that spatial cognition and the language domain can affect each other, in the way that spatial cognition can influence language (Vukovic & Shtyrov, 2017; Vukovic & Williams, 2015) but also *vice versa* where language can influence spatial processing (Majid et al., 2004). In this section, I will cover some cases of space-language interaction.

Spatial Context and Language Simulation

Language understanding necessitates the creation of a representation of the situation described in the verbal input. Furthermore, recall of such representations is required for successful recollection of what has been understood (Zwaan & Radvansky, 1998). In particular, during text comprehension readers create models that have spatial properties (Zwaan et al., 1995). For instance, Glenberg and colleagues (Glenberg et al., 1987) showed that comprehension is influenced by the spatial structure of the events described in the text and affects the foregrounding process. Hence, perceptual symbols derived from our experience during reading helps to build a situation model (Barsalou, 1999; Fincher-Kiefer, 2019; Glenberg et al., 1987). Situation models are mental representations of the state of affairs described (explicitly and implicitly) in a text (Zwaan et al., 1995; Zwaan & Radvansky, 1998). When readers comprehend each tale or text, they update their model based on five indices (time, space, causality, protagonist, and intentionality). More recently, Gianelli and colleagues (Gianelli et al., 2011) studied the impact of spatial information on ACE. They used a modified ACE task, where all the sentences were in the third-person (e.g., "*Louis gave a pen to Léa*" and *vice versa*) and consequently this type of sentence prevented the participants to take the first-person perspective (i.e., the transfer is between two external actors). As predicted, this hampered the ACE. However, this effect was restored when a spatial cue was provided (i.e., the position to adopt by the reader). Particularly, ACE effect was evident when the reader adopted the spatial position of the agent (rather than the receiver) and when was located on the right side of a scene.

Crucially situated models may provide useful information for mental (spatial and motor) simulation (Beveridge & Pickering, 2013). According to the spatial grounding hypothesis, motor simulations are grounded in the spatial context (Beveridge & Pickering, 2013). If there is minimum spatial information sufficient to grasp the spatial relationships between the arguments, we can simulate different actions other than that of the agents, including that of the receiver of an action (i.e., grammatical patient) or that of an external observer. For example, the presence of a self-referential pronoun ('I' or 'you' in the subject position, or 'me' and 'you' in other syntactic positions) should trigger the reader to assume the perspective consistent with the pronoun, centering the scene on her/his body (Beveridge & Pickering, 2013). Conversely, when no spatial cue is given, the reader cannot embody any of the actor's perspectives and the motor simulation cannot occur. This hypothesis explains some disparities in outcomes between first-person and third-person language. Papeo and co-authors (Papeo et al., 2011) discovered that when people hear a first-person sentence, they adopt an embodied agent's perspective, but not when they hear third-person sentence; however, Tomasino et al. (Tomasino et al., 2007) found that first- and third-person language generated identical action perspectives. These conflicting results are explained in the following way. The first-person sentences in Papeo's study anchor the scenario model in the comprehender's own body, permitting action simulation; the situation model in the third-person sentences lacks sufficient spatial information for action simulation. The aim of Tomasino and colleagues' (2007) study was to determine whether the described action occurred inside or outside the participant's

perspective, prompting the creation of situation models in which first- and third-person activities might be placed. According to Beveridge and Pickering (2013) the task/procedure differences might explain this incongruency. In the study by Papeo and colleagues (2011), the task did not prompt the construction of situation models (i.e., participants had to respond if the subject was in a first-person or third-person pronoun during TMS stimulation), conversely in Tomasino and co.authors (2007) the participants were required to judge whether the described action took place from an internal or external perspective, enabling the creation of models of the situation where actions are depicted.
From a behavioral point of view, readers acquire an agent's point of view in different ways, depending on the pronouns encountered during reading. That is to say that they create a situation model from a specific spatial perspective (Brunyé et al., 2009). Brunyé and co-authors (Brunyé et al., 2009), employing a sentence-picture verification task, showed that displaying on a PC screen sentences describing self-related actions ("*I am slicing the tomato*"/ "*You are slicing the tomato*") facilitated the adoption of the egocentric perspective compared to the allocentric perspective. In contrast, sentences describing non-self-related actions ("*He is slicing the tomato*") showed the opposite effect, with faster reaction times during picture verification of allocentric compared to the egocentric perspective. However, in a later study, Brunyé and co-authors partially failed to replicate this finding (Brunyé et al., 2016). Specifically, in their second experiment (Brunyé et al., 2016), the 'I' pronoun did not facilitate egocentric perspective-taking compared to the 'You' pronoun. Lastly, Vukovic and Sthyrov (Vukovic & Shtyrov, 2017) showed that hearing 'You' sentences facilitated egocentric perspective processing compared to the allocentric perspective, whereas hearing 'I' sentences displayed the opposite pattern.
Hence, task' demands or instructions may thus play a significant role in action language comprehension, in the sense that they enable or encourage participants to generate a spatial context for the stated acts.

Spatial Frames of Reference and Sentence Simulation

As described in the Embodied Space section, egocentric (body-dependent representation) and allocentric (body-independent representation) spatial frames of reference can be used to process the environment (Burgess, 2008). In addition, spatial frames of reference enable us to depict, respectively, the first-person (our point of view) and third-person (someone else's point of view) perspectives (Beveridge & Pickering, 2013; Tversky & Hard, 2009). The activation of a wide brain network during both spatial navigation and sentence-picture verification tasks supports the relationship between language and spatial frames of reference (Vukovic & Shtyrov, 2017). Vukovic and Shtyrov (2017) found that shared activity during navigation and sentence-picture verification tasks emerged in the motor, extrastriate, premotor, and anterior cingulate cortices. Motor and anterior cingulate areas are related to the egocentric frame, whereas the extrastriate cortices involve also the allocentric frame.

Crucially, individual proclivity could affect action simulation and sentence comprehension. Brunyé and colleagues (Brunyé et al., 2016) showed that the performance in their sentence-picture verification task the ability to adopt egocentric and allocentric perspectives is modulated by their propensities to be empathically involved during reading. More specifically, related to individual differences, in spatial cognition Vukovic and Williams (Vukovic & Williams, 2015) studied how spatial frames of reference proclivity (see, Gramann, 2013) influences embodiment during a sentence-picture verification task. The authors assessed spatial frame preferences through a virtual corridor navigation task, where only visual flow information was used to infer the starting position. Based on the results on that task, the participants were split into two categories: egocentric or allocentric strategy preference. Afterwards, a sentence-picture verification task was administered to assess the embodiment in egocentric or allocentric point of view after hearing first and third-person action-related sentences (see above for examples). They found that individuals with a preference towards the egocentric frame were significantly faster to verify egocentric pictures compared to allocentric pictures only for the ‘You’ sentences, whereas the allocentric preference group did not differ in answering to ‘I’ or ‘You’ sentences. They concluded that individuals with an egocentric preference simulate action-related sentences from an egocentric point of view.

In the second set of experiments, they provided a spatial context (a photo with two actors speaking; i.e., one saying to the other an action-related sentence – “*I open the bottle*”, “*You open the bottle*”). When the participant had the same egocentric perspective as the actor that was pronouncing the sentences (i.e., as if the participant was looking at his back from the photo showed on the PC screen), they found that for the ‘I’ sentences the egocentric group was faster to respond for egocentric compared to allocentric pictures, whereas the allocentric group showed the opposite trend. The authors suggested that, when providing the spatial context and the actors positions, both groups can embody the speaker’s point of view. Lastly, when the actor speaking was looking toward the participant (namely had a different point of view of that scene) the egocentric and allocentric individuals kept their preferred spatial perspective: the egocentric group the one of the speaker, and the allocentric the one of an external observer. Vukovic and Williams (Vukovic & Williams, 2015) concluded that linguistic elements, such as pronouns, are not the only determinants that affect language simulation: a crucial role is played by spatial cognition frames preferences and the presence of the spatial context (see, Beveridge & Pickering, 2013). They also suggested that the role (i.e., agent and receiver of action) of the actors in the sentence and scene could be a determinant element for action simulation.

Spatial Distance as Linguistic Metaphor

Space is often used to support linguistic metaphors as people frequently use the terms ‘close’ and ‘far apart’ to describe similar and dissimilar things (Casasanto, 2008). Lakoff and Johnson (Lakoff & Johnson, 1999) proposed two complementary conceptual mappings to describe the connection between space and similarity: spatial closeness as similarity and spatial distance as dissimilarity or alternatively, similarity as spatial closeness and dissimilarity as spatial distance. Following the conceptual metaphor theory (Lakoff & Johnson, 1980, 1999), these metaphors are grounded in our sensorimotor experience. EC theories support the fact that conceptual metaphors are grounded in our experience with the body and environment, in other words, the domains of space, force, or motion, which are experienced through our sensorimotor system, are used to describe abstract concepts and create metaphors (Barsalou, 1999, 2008; Gallese & Lakoff, 2005; M. Wilson, 2002). The importance of the space for the representation of similarity is supported by a consistent body of evidence, particularly within the semantic domain. In one important study, Casasanto (2008) presented participants with stimulus pairs (abstract nouns, faces, object pictures) at different mutual distances (far or near, far or close to each other) on a computer screen, and then they were asked to judge how much the items were similar in their meaning (abstract nouns), functional use (objects), or visual appearance (objects and faces). Findings showed that abstract nouns and object pictures (functional use) pairs were rated as more similar when they were presented in the near space (as opposed to the far space); the opposite pattern was found for face pairs and object pictures (visual appearance). The author stated that when participants were asked to make conceptual judgments (namely, when they were asked to make judgments about the meaning of nouns and the functional use of objects), the spatial proximity influenced similarity judgments; conversely, when participants were asked to make perceptual judgments, the spatial proximity operates in the opposite direction. Guerra and Knoeferle (Guerra & Knoeferle, 2014) showed that spatial proximity can affect subsequent abstract sentence comprehension of semantically related nouns. When noun pairs (e.g., ‘joy and ‘euphory’) were presented near on the PC screen and the sentence in which they were contextualized expressed similarity (“*Joy and euphory are almost similar*”), participants’ reading times at the adjective (i.e., ‘similar’) were faster compared to the condition in which the nouns were far; conversely, when noun pairs were presented far apart compared to the condition in which they were close, participants’ reading times at the adjective were faster for sentences that conveyed dissimilarity. Boot and Pecher (Boot & Pecher, 2010) showed that participants were faster to make judgments on coloured squares when they were similar in colour and near, compared to the condition when the squares were far; in addition, they were faster to respond when the squares were far and dissimilar in colour compared to near. They also found that the relation between similarity and closeness is asymmetrical, supporting the notion that similarity entails the spatial properties of closeness, but not the other way round. This is finding is complementary to the finding of Casasanto (2008) where closeness entails similarity; hence, Lakoff and Johnson's (Lakoff & Johnson, 1999) complementary conceptual mappings are both possible.

Chapter 2: The present work

Despite the discoveries in the field of space-language interaction within the framework of EC, some aspects remain unsolved or need further clarification.

Indeed, concerning the case of ‘spatial context and language comprehension, it is unclear how motor simulation and spatial perspective-taking interact when two actors (agent and receiver) are present in a sentence. The literature has focused on sentences involving only one character (Brunyé et al., 2009, 2016; Vukovic & Shtyrov, 2017; Vukovic & Williams, 2015). This issue was raised also by previous studies and hypotheses which showed that spatial information is crucial to embody the agent of a sentence (Beveridge & Pickering, 2013; Gianelli et al., 2011). In addition, it is unclear whether spatial information in situation models is constant across languages. In other words, if EC mechanisms are cross-linguistic and universal (Sinha & Kristine, 2001) or depend on the grammatical rules (Chomsky, 1980).

Regarding the case of ‘spatial frames of reference and sentence simulation’, again it has been only studied in sentences where only one actor is involved (Vukovic & Williams, 2015). However, other aspects should be addressed by future research. In the study by Vukovic and Williams (2015) spatial frame proclivity has been evaluated with a spatial orientation task (i.e., pointing to the starting position after a path; Gramann et al., 2005); however spatial cognition is sustained also by spatial memory, which is crucial to learn and retrieve new paths and item locations. A way to assess egocentric and allocentric spatial memory is through the use of VR landmark-based navigation tasks (Doeller et al., 2008; Guderian et al., 2015), where environmental features are removed to force the use of a specific frame of reference to retrieve spatial information from memory. In addition, the sentence-picture verification task forces the use of an online simulation of the sentence, but it is unclear if spatial memory performance affect offline simulation (M. Wilson, 2002) of action language.

Lastly, the tight link between space and linguistic metaphors has been observed in the domains of semantics (i.e., the meaning of words, functional use of object) and visual perception (e.g., face/object/shape appearance) during online processing (Boot & Pecher, 2010; Casasanto, 2008; Guerra & Knoeferle, 2014). It is unclear if other aspects of language and perception, such as phonology, show the same connection with space. Specifically, if the concept of spatial closeness/distance, prompted by similar/dissimilar phonology, is stored offline in the declarative memory system.

This thesis aims at elucidating at a macroscopic level the interaction between spatial cognition and language within the EC framework but has further specific objectives that will help to better understand EC mechanisms.

Four experiments will show the interaction between space and language from an embodied perspective.

- Study 1 tries to replicate the LLE through the ACE procedure.
- Study 2 wants to explore 1) the interaction between motor simulation and spatial perspective-taking in action language involving two characters - agent and receiver - and 2) if such interaction is constant in two languages to understand cross-linguistic EC processes.
- Study 3 wants to understand how egocentric and allocentric spatial memory performance affects the offline simulation of action language involving two actors (agent and receiver).
- Study 4 wants to explore if phonologically similar/dissimilar stimuli affect spatial distance estimation during memory recall.

Study 1 - LLE Replication Study

Introduction

The connection between action and language has been studied in several behavioral experiments (Andres et al., 2015; Marino et al., 2012, 2013), leading to the so-called ACE (Glenberg & Kaschak, 2002). In the ACE protocol, participants are prompted to make judgments on sentences describing actions toward the body (e.g., "*Manuela dealt the cards to you*") or away from the body (e.g., "*You dealt the cards to Manuela*"); crucially to respond they could perform either a movement of the hand/arm away or toward the body. The mismatch between the action executed and the action implied by the sentence yields an increase in the reaction times needed to perform the task (Glenberg & Kaschak, 2002). This result is usually explained in terms of embodied simulation: when presented with the sentence, participants simulate its content, therefore reactivating the motor pattern described; if the motor pattern re-instantiated during simulation is incompatible with the action planned to respond, the mismatch between the two creates an interference effect.

For years, the ACE was studied and confirmed by independent research groups. It was replicated in different languages [e.g., Italian (Glenberg et al., 2008), Spanish (De Vega et al., 2013), and Japanese (Shunji, 2011)] and at different timings of the motor response (Borreggine & Kaschak, 2006; Diefenbach et al., 2013; Kaschak & Borreggine, 2008). Moreover, the paradigm was extended to include rotatory movements (Zwaan & Taylor, 2006) and body posture (Zwaan et al., 2012). Beyond behavioral pieces of evidence, neural correlates of ACE were also found: Aravena and collaborators (Aravena et al., 2010) recorded event-related potentials while participants underwent an ACE task and compared compatible *vs.* incompatible conditions. The authors observed an N400-like component localized in Cz and associated with the incompatible condition, whereas a motor potential and a re-afferent potential were observed during trials in the compatible condition. Taken together, all these studies consistently reported the presence of the ACE under the described experimental conditions. Notwithstanding, other studies failed to find the ACE. The first study questioning the reliability of the ACE was by Papesh (Papesh, 2015) who did not replicate the results in any of the 8 experiments they carried out for this purpose. After running Bayesian analyses, the authors discussed these findings suggesting the possibility that in previous studies the effects reported as significative might rather be compatible with the null hypothesis.

Recently, Diez-Alamo and colleagues (Díez-Álamo et al., 2020) confirmed the unreliability of ACE in experiments testing not only the effect in the standard linguistic task (as in the original paradigm) but also in the incidental memorization of the sentences. In their work, the ACE failed to emerge, but eventually, another pattern of results appeared that was consistent across the 5 experiments reported. Specifically, individuals were faster in understanding, and they memorized better the sentences that described an action toward the body. This effect has been referred to as LLE. The name is inspired by perceptual research, where the looming effect is related to the overestimation of sound intensity or object's speed when they approach the listener/viewer (Neuhoff, 2001, 2018). Different explanations of the LLE have been put forward. The first, in continuity with the hypotheses, suggested to account for the twin perceptual effect, refers to its evolutionary and adaptive value: it is conceivable that objects approaching us are more relevant to process since they could be either desirable or harmful. Therefore, aiming at guaranteeing survival, a fast evaluation is needed to coherently adapt our behavior (acceptance *vs*. withdrawal). In this view, the linguistic counterpart (LLE) could be the simulation of a real perceptive bias (Glenberg & Gallese, 2012). Furthermore, the authors suggested that LLE is coherent with other frameworks, such as the mental models' theory of reading (Gunraj et al., 2014; Zwaan & Radvansky, 1998). According to this theory, as we read a sentence, we build a mental model of the content that includes all the elements and the actions described in the sentence. Later on, we are faster and more efficient in processing those elements of the model that were related to the participant (Glenberg et al., 1987). Indeed, the toward sentences refer to something that is getting closer to the subject. Another possible interpretation is related to the well-known advantage of self-referenced information for memorization, according to which we tend to recall better information about ourselves (Sui &

Humphreys, 2015). In this perspective, toward sentences would be better retained because they describe something related to the subject.

The present study originated from this line of research. In two cross-linguistic studies (Italian and U.S. English), I tried to replicate the LLE. Consequently, I predict that ACE will not emerge and instead a main effect of sentence direction in favor of toward phrases (LLE). In addition, I want to test these hypotheses in two languages (U.S. English and Italian); consequently, I expect that the predictions hold in both samples.

Methods (U.S. English Experiment)

Participants

Thirty-eight younger-adult undergraduate students were recruited from the University of Hartford subject pool and 35 were included in the statistical analyses (M_{age} = 18.76, SD_{age} = 1.34; 12 males, 33 right-handed). All participants reported no current history of serious psychiatric or neurological disorder as assessed through self-report. Most subjects were freshmen students (M = 12.6 years of education) and all participants had normal or correct-to-normal vision. All participants gave written consent as approved by the Institutional Review Board of the University of Hartford and participated in the study for course credit.

Materials

Stimuli presentation and participant response were carried out with Eprime 2.0 and presented on a Dell Optiplex 7020 computer. The experiment consisted of 160 total trials composed of 40 abstract sentences (20 away, 20 toward), 40 concrete sentences (20 away, 20 toward), and 80 nonsense sentences were presented in English to participants. All the sentences were structured in a subject, verb, direct object, indirect object format with the sentence subject either a third-person actor or the second-person "you" (e.g., "*Andy pitched the idea to you*" and "*You kicked the soccer ball to Joe*", respectively). All sentences were presented in the simple past tense (see Appendix A). The abstract, concrete, and nonsense sentences were pseudo-randomly presented with brief break periods after every 40 trials presented. Responses were made on a modified keyboard specifically designed for the task, with all keys removed except for three response keys: the q, p, and 7 (keypad) keys. The q and 7 keys were color-coded as blue and green, and the center p key was colored as white. All other keys were removed and the keyboard itself was wrapped in black duct tape. The experiment started with green as indicating a nonsense sentence and blue as sensible, and then at the halfway point (i.e., after 80 test trials), flipping the keyboard 180º although the color-coding remained the same.

Procedure

The experiment was conducted in a quiet 10' x 15' private testing room. The research assistant who oversaw the test session quietly sat in a corner of the room away from the line of sight of the subject. For the experiment, participants were seated at a distance of 60cm from the viewing screen. The keyboard was placed orthogonally between the computer screen and the participant. Instructions were presented both on screen and verbally by the research assistant: *"In this task, you will be shown a sentence on the screen. Your job will be to read the sentence to yourself. Your job is to decide whether the sentence makes sense or is nonsense."* Examples were provided and then instructions were given on the response keys. *"To start the trial, press and hold the white button. The sentence will appear as you hold down the white button. After you have read the sentence, decide if it is sensible or nonsense. When ready, let go of the white button and make your decision"*. Participants were then instructed on the "sensible" or "nonsense" response buttons, with a paper reminder sheet displaying the coding placed alongside the computer for a reminder. Participants were specifically instructed to respond using the same index finger for each trial and were reminded to respond as quickly and as accurately as possible. The testing time for the experiment was around 20 minutes and around 60 minutes for the entire testing session.

Statistical analyses

All analyses presented in this paper were performed by using R (R Core Team, 2014), version 3.6.3. Linear mixed-effects [*lme4* package (Bates et al., 2015)] ANOVAs were carried out with Kenward-Roger approximation and restricted maximum likelihood estimation (Luke, 2017). Single term deletion was used to determine the significance of random effects (RE) in the model (Bates et al., 2015). All RE were set as having random intercepts, as all models failed to converge when allowing for random intercept and slope for these effects. The R formular was the following: [lmer(RT ~ response condition*sentence direction*sentence type + (1|participant)]. The dependent variable was the sentence comprehension RT and the fixed effects were response condition, sentence direction, and sentence type. Variance explained by RE was calculated as ((σ^2 RE$_1$ + σ^2 RE$_n$ /σ^2 tot) x 100). The mixed-effects model diagnostic was assured for all models by visually checking residuals distribution and homoscedasticity. *Emmeans* package (Lenth, 2018) was used to run post-hoc comparisons with Bonferroni correction. Effect size (η^2_p) was interpreted according to Richardson (Richardson, 2011) (small = 0.01, medium = 0.06 and large = 0.14). A p-value less or equal than 0.05 is considered statistically significant.

Results (U.S. English Experiment)

Three participants were removed from the analyses using the interquartile range method [IQR (1.5)], which allows identifying outliers lying below the 25th or above the 75th percentile. For each subject trial, the mean within stimuli was computed. Nlog (natural logarithm) transform was not applied to the response variable as filtering resulted in proper distribution, however, results are reported as transformed to easily compare them with Italian experiment values. Nonsense sentences, as well as omissions and errors, were removed from the analyses.

Mean participants and items accuracies were 93.30% and 93.63% respectively. To account for inter-individual variability, American participants were assigned as a random effect, which was found to be significant ($p < .001$) and explained 69.12% of the variance (namely intraclass correlation coefficient). There were three fixed effects variables: 1) Response Condition (far-near), 2) Sentence Direction (toward-away), and 3) Sentence Type (concrete-abstract), resulting in a 2x2x2 within-subjects ANOVA. Response Condition by Sentence Direction by Sentence Type interaction (i.e., ACE) was not significant ($F_{1, 238} = 1.15$, $p = 0.284$, $\eta^2_p = 0$, 95%CI[0, 0.04]).

Results showed a significant main effect of Sentence Direction ($F_{1, 238} = 22.36$, $p < 0.001$, η^2_p = 0.09, 95%CI[0.02,0.14]), where away sentences yielded faster comprehension times ($M_{log(away)}$ = 7.58, $SE_{log(away)}$ = 4.27) than toward sentences ($M_{log(toward)}$ = 7.62, $SE_{log(toward)}$ = 4.27). This is opposite to the LLE. A significant Sentence Direction by Sentence Type interaction was found ($F_{1, 238} = 5.75$, $p = 0.017$, $\eta^2_p = 0.02$, 95%CI[0, 0.07]), where only for concrete sentences ($p < 0.001$) the away responses were faster than toward ones (abstract, $p = 0.1$). Moreover, a significant main effect of Response Condition ($F_{1, 238} = 57.77$, $p < 0.001$, $\eta^2_p = 0.20$, 95%CI[0.11, 0.28]) showed that responses in the near condition were faster than in the far condition. Again, a significant Sentence Type by Response Condition interaction ($F_{1, 238} = 18.42$, $p < 0.001$, $\eta^2_p = 0.07$, 95%CI[0.02, 0.14]) was found; near responses were faster especially for abstract ($p < 0.001$) than concrete ($p = 0.201$) sentences. No other effect was significant. According to these findings, I failed to find the ACE, but also I did not replicate the LLE. Figure 1 (left panel) shows the significant results in this sample.

Methods (Italian Experiment)

Participants

Twenty-one young adults ($M_{age} = 25.70$, $SD_{age} = 7.46$; 7 males; 19 right-handed; 16 with master's degree, 4 with bachelor's degree, 1 with Italian high-school degree) were recruited at the Catholic University of Milan. All participants were assessed through a brief anamnestic questionnaire by a licensed psychologist, concerning the current use of medications and recent history of psychiatric or neurological disorders. All participants had normal or correct-to-normal vision. A minimum sample size of 22 was determined by using a Cohen's *d* of 0.64 (as derived from the main effect of sentence direction in U.S. English experiment above) and computed with a power of 0.8 and a total of 40 stimuli by using an online linear mixed-effects (CNC design) model calculator (Westfall et al., 2014). Participants gave written consent as approved by the Ethical Committee of the Catholic University of Milan and participated in the study without any compensation.

Materials

Forty concrete sentences (20 away, 20 toward) and 40 nonsense sentences were taken from the stimuli of U.S. English experiment, translated, and adapted into Italian (see Appendix B). All the sentences were modified, replacing the second-person pronoun 'you' with the first-person pronoun 'I', in the correct thematic role according to the syntactic structure. This was done to better promote the embodied motor simulations of self-other (first-person *vs.* third-person) distinction (Schütz-Bosbach et al., 2006), and to promote the inclusion of the self into the mental model (Zwaan & Radvansky, 1998). Besides, I changed the verb tense from the past to the present continuous, as one study has underlined that the grammatical aspect can modulate how the simulations are performed (Bergen & Wheeler, 2010).

Procedure

Participants were welcomed in a quiet room with constant light conditions and were given the consent form. Once signed, demographic (age, sex, education, dominant hand) and brief anamnestic information (history of neurological, psychiatric disorders, and relevant conditions or medications) were gathered by a licensed psychologist. Participants were seated at a distance of 50cm from a 15'' laptop. A keyboard was placed orthogonally between the screen and the participant. G button was masked by a blue layer, A (far button), and L (near button) by yellow and red layers respectively. Before each task block, a practice block of 12 trials was provided. Instructions were "*Use your right index finger to answer. Once a + appears, hold down the BLUE button and read the sentence until you understand the meaning. Press the YELLOW button if the sentence makes sense, if it does not make sense press the RED button. To go to the next sentence, press and hold BLUE.*" For the near response, only this part was changed "*Press the YELLOW button if the sentence does not make sense, if it makes sense press the RED button.*" The order of blocks, with either 'yes-is-near' or 'yes-is-far' instructions, was counterbalanced across participants. A debriefing was provided at the end of the task.

Statistical analyses

The statistical approach was the same as that of the U.S. English experiment. The R formular was the following: [lmer(RT ~ response condition*sentence direction + (1|participant)+ (1|stimuli)]. The dependent variable was the sentence comprehension RT and the fixed effects were response condition and sentence direction. In addition to participants, I added the sentences as a random effect.

Results (Italian Experiment)

Reaction times between 100ms and 5000ms were retained for each trial, whereas error and miss responses were removed from analyses. To improve the distribution of the response variable, I applied IQR (1.5) method for outliers' detection for trials within each participant, condition (far-near), and sentence direction (toward-away) (trial removed = 94/1688; imputed with corresponding individual's mean for condition and sentence direction); however, this was not sufficient, and I then Nlog-transformed the dependent variable to ameliorate skewness (values between +/-1). Nonsense sentences, as well as omissions and errors, were removed from the analyses.

The average total accuracy for participants and stimuli were 95.48% and 98.49%. In this study, Italian participants were included as random effects in order to account for inter-individual variability and, in addition to Experiment 1, I checked inter-stimuli variability in the model. The model did not to converge due to the relatively small sample size compared to complex random effect structure (variance for stimuli was 3.7%). Hence, I included only participants as random effect in the model and convergence was met. This random effect was found to be significant ($p < 0.001$) and explained 53.54% of the variance. Fixed effects were the Response Condition (far-near) and Sentence Direction (toward-away), resulting in a 2x2 within-subjects ANOVA. As for the USA sample, Response Condition by Sentence Direction interaction (i.e., ACE) was not significant ($F_{1, 1546} = 0.27$, $p = 0.600$, $\eta^2_p = 0$, 95%CI[0, 0]). Again, I found a significant main effect of Sentence Direction ($F_{1, 1546} = 74.76$, $p < 0.001$, $\eta^2_p = 0.05$, 95%CI[0.03, 0.07]): away sentences were understood faster ($M_{log(away)}$ = 7.31, $SE_{log(away)}$ = 0.05) than toward sentences ($M_{log(toward)}$ = 7.41, $SE_{log(toward)}$ = 0.05). This is again in contrast to the LLE. No other significant effect was found. In line with the findings of the U.S. English experiment, I found neither the ACE nor the LLE. Figure 1 (right side) shows the main effect of sentence direction for the Italian sample.

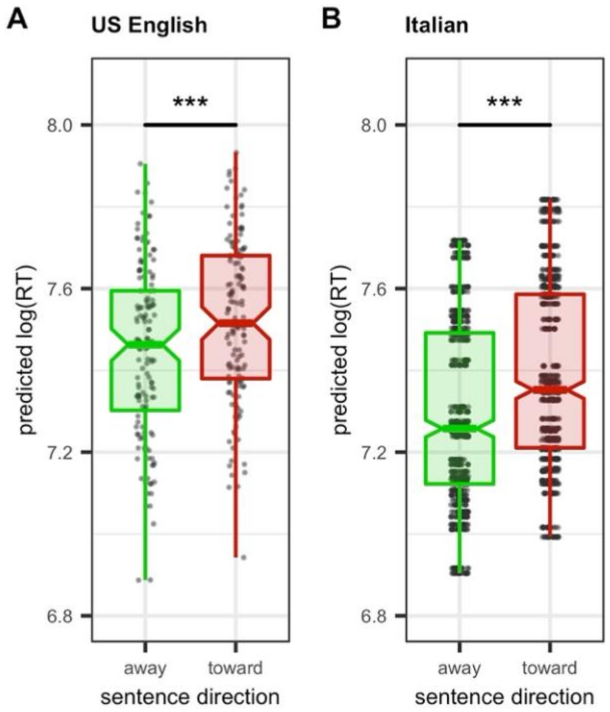

Figure 1 The effect found in the two samples ($p < 0.001$). (A) U.S. English experiment (black dots represent each individual mean for the stimuli in the conditions). (B) Italian experiment (black dots represent every single trial for each individual in the conditions). *** $p < .001$

General Discussion

In this two experiments I sought to replicate the LLE and faster processing times for toward compared to away sentences (Díez-Álamo et al., 2020).

In accordance with the predictions, I failed to replicate ACE in both languages but surprisingly, I found an opposite direction in the U.S. English and Italian sample. Away sentences were processed (i.e., reading times) faster compared to the toward ones. This was true even when slight changes in the methodology and stimuli (‘I’ vs. ‘You’ first-pronoun sentences) were applied in the two experiments conducted in USA and Italy. I called this the *Linguistic Inverse Looming Effect* (LILE). Interestingly, I suggest that the LILE could connected to the agent’s linguistic perspective rather than the described action, where sentences in first-person (‘I’ or ‘You’) were processed faster than those in the third-person perspective (‘He’/’She’). Such difficulty to replicate EC behavioral studies might depend on procedural, methodological, and inter/intra individual differences (Beveridge & Pickering, 2013; Vukovic & Williams, 2015).

According to previous findings (Beveridge & Pickering, 2013) investigating the reader’s point of view while processing a sentence, different pronouns could induce different linguistic perspectives: in particular, the first-person and second-person pronouns in the thematic role of the participant in the position of the agent could trigger the simulation of the action described from this perspective (Brunyé et al., 2009). Following this line of reasoning, I hypothesize that the LILE might be triggered by the spatial perspective assumed by the reader. Such explanation is not in contrast but complementary to the LLE (Díez-Álamo et al., 2020), as embodiment might depend on several factors, such as individual preferences, task demands, and situation models created by the subject and task instructions.

However, this study has some limitations. First of all, the two studies are not comparable in terms of methodology and stimuli adopted, thus it is hard to demonstrate a clear cross-linguistic effect; than samples are small and from only two laboratories.

The next study aims at ameliorating the stimuli and methodology in the two samples and to explore the complex interaction between motor simulation and spatial perspective-taking when two characters are involved in a sentence.

Study 2 - Spatial Context and Sentence Simulation

Introduction

As we read sentences or a text we create situation models that have spatial properties (Zwaan et al., 1995; Zwaan & Radvansky, 1998). Furthermore, this situation models can be depicted from different perspectives (Beveridge & Pickering, 2013). The idea that we can assume multiple perspectives in language understanding is consistent with the extant literature on spatial cognition. Indeed the environment can be represented with two spatial frames of reference (Burgess, 2006, 2008): egocentric frame (body-dependent representation) and allocentric frame (body-independent representation). These allow us to represent respectively the first-person perspective (our point of view) and third-person perspective (someone else's point of view) (Beveridge & Pickering, 2013; Tversky & Hard, 2009).

Behavioral studies support the hypotheses that we create situation models with different spatial perspectives. Brunyé and colleagues (Brunyé et al., 2009), using a sentence-picture verification task, found that showing on a PC screen sentences describing self-related actions ("*I am slicing the tomato*"/"*You are slicing the tomato*") facilitated the adoption of the first-person perspective (egocentric) compared to third-person perspective (allocentric) when matching the sentences to photos. In contrast, sentences describing non-self-related actions ("*He is slicing the tomato*") showed the opposite effect, with faster reaction times for matching the sentence to pictures representing an allocentric point of view of the action. However, these results are questionable as Brunyé and co-authors partially failed to replicate their previous findings (Brunyé et al., 2016). Specifically, in their second experiment (Brunyé et al., 2016), they demonstrated that the 'I' pronoun did not facilitate egocentric perspective-taking compared to the 'You' pronoun. Moreover, Vukovic and Sthyrov (2017) showed that hearing 'You' sentences facilitated first-person perspective photo processing compared to third-person perspective photo, whereas hearing 'I' sentences displayed the opposite pattern.

Uncertain conclusions can be drawn from these results. On the one hand, the process of understanding complex sentences calls into play both motor simulation and spatial perspective-taking, with different degrees of involvement depending on the task, the syntactic structure, and the specific content of the sentence. On the other hand, the literature is currently unclear about if and how the processes of motor simulation and spatial perspective-taking interact in sentences with two characters (i.e., the participant and someone else, see also ACE results) in which spatial cues (e.g., a photo of the action represented in the sentence) are provided. In the current set of experiments, I applied a two-agents sentence-picture verification task I created using a procedure similar to that of Brunyè and colleagues (2009, 2016) and evaluated it in two languages (Italian and U.S. English). Crucially, and differently from previous studies, I manipulated agency (who is acting) both in the sentence and in the photo. Four experimental conditions were created: two congruent conditions and two incongruent conditions. In the first congruent condition, the participant is the agent in the sentence (i.e., "*'I' am giving John the pen*") and photo (i.e., the photo depicts from an egocentric point of view the hands of the participant acting). In the second congruent condition, the agent is a third person (i.e., "*John is giving 'me' the pen*") and in the photo the agent is someone else in front of the participant (i.e., the photo depicts the participant's egocentric point of view on the hands of someone else acting). In the incongruent conditions, there is an incongruency between the agent in the sentence and that in the image. According to the spatial grounding hypothesis (Beveridge & Pickering, 2013), the presence of the self-referential pronoun should prompt readers to assume the corresponding perspective, and as a consequence, they should simulate the action from that perspective. Therefore, I expect that if motor simulation and spatial perspective-taking always match, the two congruent conditions should yield similar verification times since in both cases the photos correctly represent the perspective and the direction of the movement described in the sentence. A converse possibility may arise if the strong view of embodiment is correct. According to this view (Decetey, 2002), I always perform motor simulations as the agent of the action, even when the spatial perspective I assume is not that of the agent. As such, the condition in which I am the agent (with the pronoun 'I' as a subject) should yield faster reaction times compared to the condition in which the first-person pronoun is in the thematic role of the receiver (with the pronoun 'me'). In this latter case, motor simulation and spatial perspective would not overlap, and the picture would match only in the perspective of the action but not in its direction.

Lastly, the experiment's cross-linguistic design was carefully planned, including distinct cohorts of Italian and Northern American volunteers. Cross-linguistic research helps establishing potentially universal trends while adjusting for region-specific linguistic variations. I predict that these mechanisms are cross-linguistic and are present irrespective of languages (in this study Italian and US English), thus reflecting common embodied mechanisms.

Methods (Italian Experiment)

Participants

Sixty-eight Italian-speaker adults (M_{age} = 26.66, SD_{age} = 6.95; 36 males; 61 right-handed; 14 with master's degree, 18 with bachelor's degree, 32 with Italian high-school degree, 2 with high specialization master and 2 with technical degrees) were recruited through Prolific (https://www.prolific.co/). Online recruitment was carried out during the COVID-19 pandemic (July-October 2020). Selection parameters of participants in the platform were set according to the self-reported absence of a history of psychiatric or neurological disorders (including language disorders), use of psychotropic drugs, and normal or corrected-to-normal vision. Participants were paid 2.13€ for performing the experiment, which lasted 15 minutes. To determine the sample size, I used the effect size of a previous study that used a similar task (Brunyé et al., 2016). With a Cohen's *d* of 0.35, a power of 0.9, and 80 stimuli, the power analysis for a mixed-effects model (Westfall et al., 2014) required a minimum of 63 participants. Participants were recruited online and gave their consent to participate as approved by the Ethical Committees of the Catholic University of Milan.

Sentence-picture verification task

A new version of the sentence-picture verification task from Brunyé and co-authors (Brunyé et al., 2009) was created with Gorilla (Anwyl-Irvine et al., 2020), an online experiment builder for cognitive tasks. Forty sentences (20 in first-person e.g., "*I am passing the tray to Marc*", and 20 in third-person e.g., "*Paul is passing me the tray*", with the agent respectively being the first-person pronoun or a third-person subject) similar to Glenberg and Kaschak stimuli (Glenberg & Kaschak, 2002) were used (see Appendix C). In particular, sentences were modified by replacing the 'you' with the 'I' pronoun and used in the present tense form as in the original task by Brunyé and colleagues (2009). Forty wrists-to-hands photos depicting the 40 sentences were taken (accessories were removed, e.g., clocks or rings). The actor in the photo was right-handed and performed the actions always with the right hand. Six similar sentences and related photos were added to create the practice trials. Twenty photos represented actions from the camera's point of view (as if the viewer was acting), while in the other 20 images the same actions had the opposite direction (someone else in front of the camera was acting). The camera was placed either over the front of the actor with an elastic band or on a tripod in front of the actor. The stimuli were validated to see whether they portrayed well the sentences depicted. Fifty-four adults (age range 23-68) participated in the validation experiment (46 included after removing individuals with vision problems, self-reported language disorders, or spatial disorientation episodes). The sentences were presented one by one followed by the congruent image. Participants were asked to rate the extent to which the image was depicting the action of the sentence ("*Does the photo match the sentence just shown?";* 0 = not at all - 10 = totally). Images with a total median score under five were shot again to improve hand position and/or gesture. All the photos are available at https://osf.io/7s94v/.

For the sentence-picture verification task, each sentence was paired with both the congruent and the incongruent photo. Consequently, in the congruent condition the agent was the same in the sentence and the photo, whereas in the incongruent pairs the agent was not the same in the sentence and the photo. This resulted in a list of 80 stimuli pairs with four sentence-photo pairs conditions (20 trials for each condition): 1_1 (the agent in the sentence and photo is the participant), 1_3 (the agent in the sentence is the participant but the agent in the photo is someone else in front of the subject), 3_3 (the agent in the sentence is a third person and in the photo (s)he was someone else in front of the participant), and 3_1 (the agent in the sentence is a third person but the agent in the photo is the participant). Following this design, conditions 1_1 and 3_3 are considered congruent since the agent in the sentence and picture matched, whereas conditions 1_3 and 3_1 are considered incongruent since the agent in the sentence and picture did not match. Twelve additional pairs of stimuli, not included in the main task, were used as training trials. Figure 2 shows the examples of each Sentence-Photo Pairs condition.

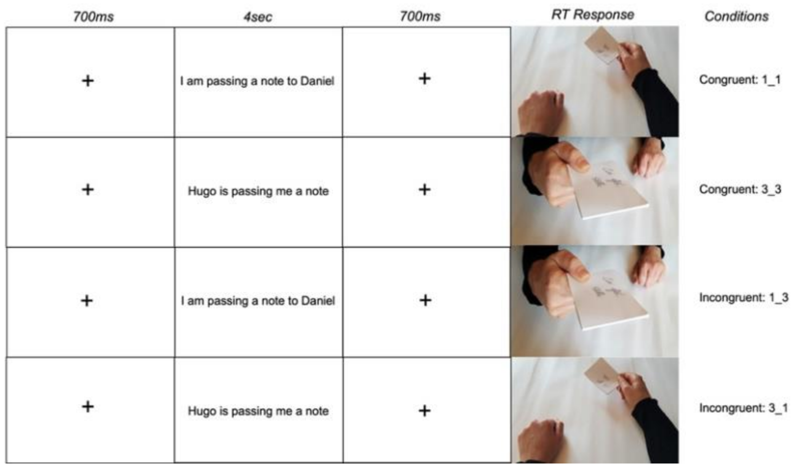

Figure 2 Example of congruent and incongruent conditions used in Experiment 1 and Experiment 2. 1_1: the reader (i.e., participant) is the agent in both sentence and photo; 3_3: the agent in the sentence and photo is someone else interacting with the reader; 1_3: the reader is the agent of the sentence but occupies the role of receiver of the action in the photo; 3_1: the agent in the sentence is a third person but the agent in the photo is the participant; RT: reaction time

Procedure

Participants who met the experiment's predetermined inclusion/exclusion criteria could access the Gorilla link to begin the task after the experiment was posted online on the Prolific system. After ticking the consent form (mandatory for proceeding further), demographical (age, sex, education, dominant hand) information was collected. First, participants did a practice block of 12 trials with the instructions (see below) with feedback for correct (green checkmark) and incorrect (red cross) responses. Participants could then comprehend how to accurately match the words to the image (i.e., by relying on the sentence-photo agent match-mismatch).

For the actual task, instructions were as follows: "*You will see a series of sentences, each followed by an image. Your job is to understand the sentence and decide if the image correctly represents the sentence you just read. To answer, place the index of the right hand on the L key and the index of the left hand on the A key from the beginning of the experiment. Press A to indicate «YES – the image represents the sentence correctly» or Press L to indicate «NO – the image DOESN'T represent the sentence correctly». Speed matters – respond as quickly as you can while still being accurate.*" Button order was counterbalanced across the participants. Each trial started with a fixation cross presented for 700ms, then the sentence was shown for 4s, followed by another 700ms-cross and the photo (response times were registered here). The stimuli pairs (i.e., sentence-photo pairs) list was randomized for each participant. The software registered reaction times and accurate responses. Four attentional checks (i.e., find and click on the cat photo among eight images of dogs) across the 80 trials were placed. Only individuals who successfully completed all attentional checks were taken into consideration (no one excluded following this criterion).

Statistical Analyses

The study is a 2X2 within-subjects experiment, with one variable being Sentence Agent (first *vs.* third person) and the second variable being the Photo Agent (the participant *vs.* someone else). The combination of the two consisted in four conditions (Sentence-Photo Pairs), labeled as follows: 1_1, 1_3, 3_1, 3_3. The first numbers represent the agent of the sentence (i.e., first or third) and the last numbers represent the agent of the photo (i.e., the participant or someone else).

All the analyses presented in this paper were performed by using R (R Core Team, 2014), version 3.6.3. Linear mixed-effects [*lme4* package (Bates et al., 2015)] ANOVAs were carried out with Kenward-Roger approximation and restricted maximum likelihood estimation (Luke, 2017). Single term deletion was used to determine the significance of random effects (RE) in the model (Bates et al., 2015). All RE were set as having random intercepts, this because all models failed to converge when allowing for random intercept and slope for these effects. Variance explained by RE on the dependent variable (picture verification reaction times) was shown with the intraclass correlation coefficient (ICC). The following formula was used in the R code: [picture verification RT ~ fixed effect + (1|participant ID) + (1|stimuli pairs ID)]. The stimuli pairs ID is a unique identifier for each (i.e., 80) sentence-image combination that the participants saw in the task. The mixed-effects model diagnostic was assured for all models by visually checking residuals distribution and homoscedasticity. *Emmeans* package (Lenth, 2018) was used to analyze post-hoc contrasts (1_1 *vs*. 3_3, 1_3 *vs*. 3_1, 1_1 *vs*. 1_3, and 3_3 *vs*. 3_1) with Bonferroni correction. To compute accuracy in among these contrasts, I calculated d-prime (d′) with the following formula [$d'(contrast_1, contrast_2) = z(p(Correct|contrast_2)) - z(p(Correct|contrast_1))$].

Benjamin and Berger (Benjamin & Berger, 2019) recommendations for improving p-value interpretation with Bayesian parameters were followed. They state that providing Bayes parameters helps to answer the question "*how strongly does the evidence favor the alternative hypothesis relative to the null hypothesis?*" which cannot be answered with the only use of p-value. Effect size (η^2_p) was interpreted according to Richardson (2011) (small = 0.01, medium = 0.06 and large = 0.14), whereas Cohen's d was interpreted according to Cohen's rule of thumb (Sullivan & Feinn, 2012) (small = 0.2, medium = 0.5 and large = 0.8). The response variable in all the studies is always reported from the predicted values of the linear mixed-effects model. α level was set to .05.

Results (Italian Experiment)

Participants with an overall trial accuracy greater than 80% were retained (4/68 removed; 64 participants included in the analyses). Similarly, stimuli with accuracy lower than 80% were removed (0 stimuli removed). Then RTs between 100ms and 3000ms were included, responses greater than 3000ms were removed as in Brunyé and co-authors (Brunyé et al., 2009). Only correct responses were included in the analyses. Natural log transformation was sufficient to improve skewness without outlier detection.

Random and Fixed Effects

The average total accuracy for the included (N = 64) participants was 96.37% and the average total accuracy for the stimuli pairs (sentence and photo) was 94.49%. 4838 observations were included in the analyses (1_1 = 1238; 3_3 = 1213; 1_3 = 1198; 3_1 = 1189) after removing incorrect responses and reaction times filtering.

In the first block of analysis, participants and stimuli pairs were put as RE and the variables Sentence Agent and Photo Agent as fixed effects (2x2 levels within-subjects design). Both RE were found to be significant ($p < .001$) and represented 30% of the variance (participant ICC = 29.4%, stimuli ICC = 0.9%) in the dependent variable (i.e., picture verification). This indicates that most of the variance was due to intra-individual variability.

Findings indicated a significant interaction (Photo Agent by Sentence Agent) ($F_{1, 76} = 60.05$, $p < 0.001$, $\eta^2_p = 0.44$, 95%CI[0.28, 0.57]). Main effect of Photo Agent was not significant, although Sentence Agent showed a statistical tendency ($p = 0.058$). Results indicated that, when the participant was the agent in the sentence, image verification times were faster when the participant was the agent in the photo ($M_{log} = 6.74$, $SE_{log} = 0.03$) than when the participant was the observer of someone else acting ($M_{log} = 6.85$, $SE_{log} = 0.03$). Conversely, when the agent in the sentence presented before the image was someone else, verification times were faster for images in which the agent was someone else ($M_{log} = 6.78$, $SE_{log} = 0.03$) than those in which the agent was the participant ($M_{log} = 6.85$, $SE_{log} = 0.03$). See Figure 3 for this result (right side).

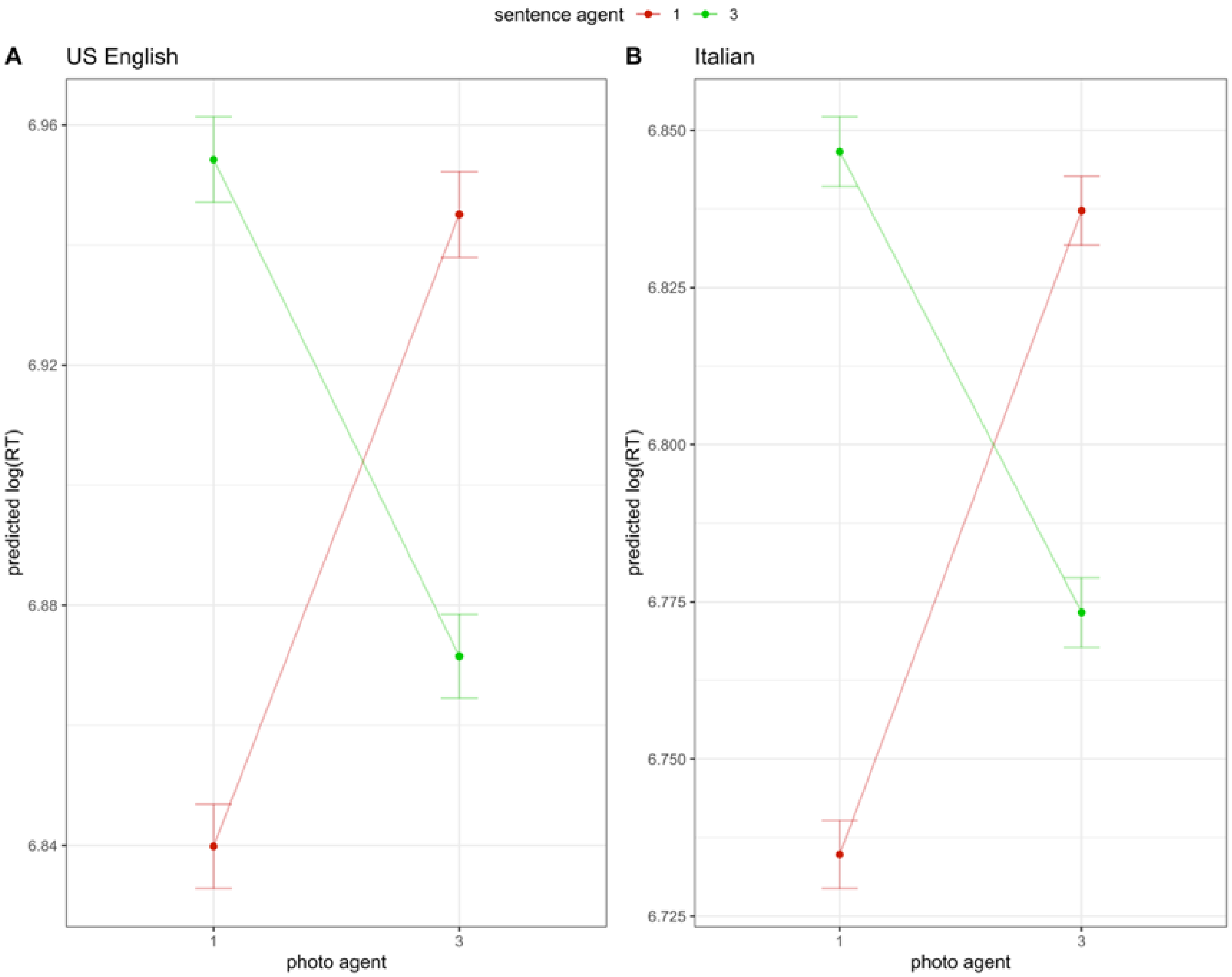

Figure 3 (A) US English sample: interaction between the agency in the photo and sentence (p < .001). (B) Italian sample: interaction between the agency in the photo and sentence (p < .001). Photo agent label 1: the participant is the agent; Photo agent label 3: the agent is someone else in front of the participant; Sentence agent label 1: the participant (i.e., reader) is the agent in the sentence; Sentence agent label 3: the agent in the sentence is a third-person interacting with the reader, who is the receiver of the action). Mean and standard errors are depicted in each graph.

As the Photo Agent by Sentence Agent interaction was significant, I ran the second block of analyses. Participants and stimuli pairs were put as RE with random intercept and the variable Sentence-Photo Pairs as a fixed effect (4 levels within-subjects design). Both RE were found to be significant (p < .001) and represented 30.33% (ICC) of the variance in the RT (participants ICC = 29.4%, stimuli ICC = 0.93%). Again, most of the RE variability on the dependent variable was due to participants rather than stimuli pairs.

Results indicated a significant effect of the Sentence-Photo Pairs ($F_{3, 76} = 21.94$, $p < 0.001$, η^2_p = 0.46, 95%CI[0.29, 0.58]). To control a potential main effect of the participant handedness, I also added this variable in the model (the main effect of this covariate and its interaction with Sentence-Photo Pairs were not significant), the final model was the one with only Sentence-Photo Pairs as a fixed effect.

As hypothesized, planned contrast showed a significant difference between the congruent conditions 1_1 and 3_3 ($t_{74.8} = -2.34$, $p = 0.022$, $d = -0.27$, 95%CI[-0.50, -0.04]). Specifically, when the agent was the participant in both the sentence and image verification times were faster ($M_{log(1\_1)} = 6.74$, $SE_{log(1\_1)} = 0.03$) than when the agent was someone else in the sentence and photo ($M_{log(3\_3)} = 6.78$, $SE_{log(3\_3)} = 0.03$). Moreover, I found significant differences between the congruent and incongruent conditions. In particular, contrasts 1_1 *vs.* 1_3 ($t_{75.5} = -6.46$, $p < 0.001$, $d = -0.74$, 95%CI[-1.00, -0.49]) and 3_3 *vs.* 3_1 ($t_{76.6} = -4.51$, $p < 0.001$, $d = -0.52$, 95%CI[-0.75, -0.28]). In the congruent condition 1_1, verification times were faster ($M_{log(1\_1)} = 6.74$, $SE_{log(1\_1)} = 0.03$) than in the incongruent condition 1_3 ($M_{log(1\_3)} = 6.85$, $SE_{log(1\_3)} = 0.03$); similarly in the congruent condition 3_3, verification times were faster ($M_{log(3\_3)} = 6.78$, $SE_{log(3\_3)} = 0.03$) than in the incongruent condition 3_1 ($M_{log(3\_1)} = 6.85$, $SE_{log(3\_1)} = 0.03$). There were no significative differences between the incongruent condition 1_3 *vs*. 3_1, suggesting that mismatching conditions have a comparable effect on picture verification. Concerning accuracy, no differences in d-prime were found among each contrast. Figure 4 (right side) represents the effects found in the Italian sample.

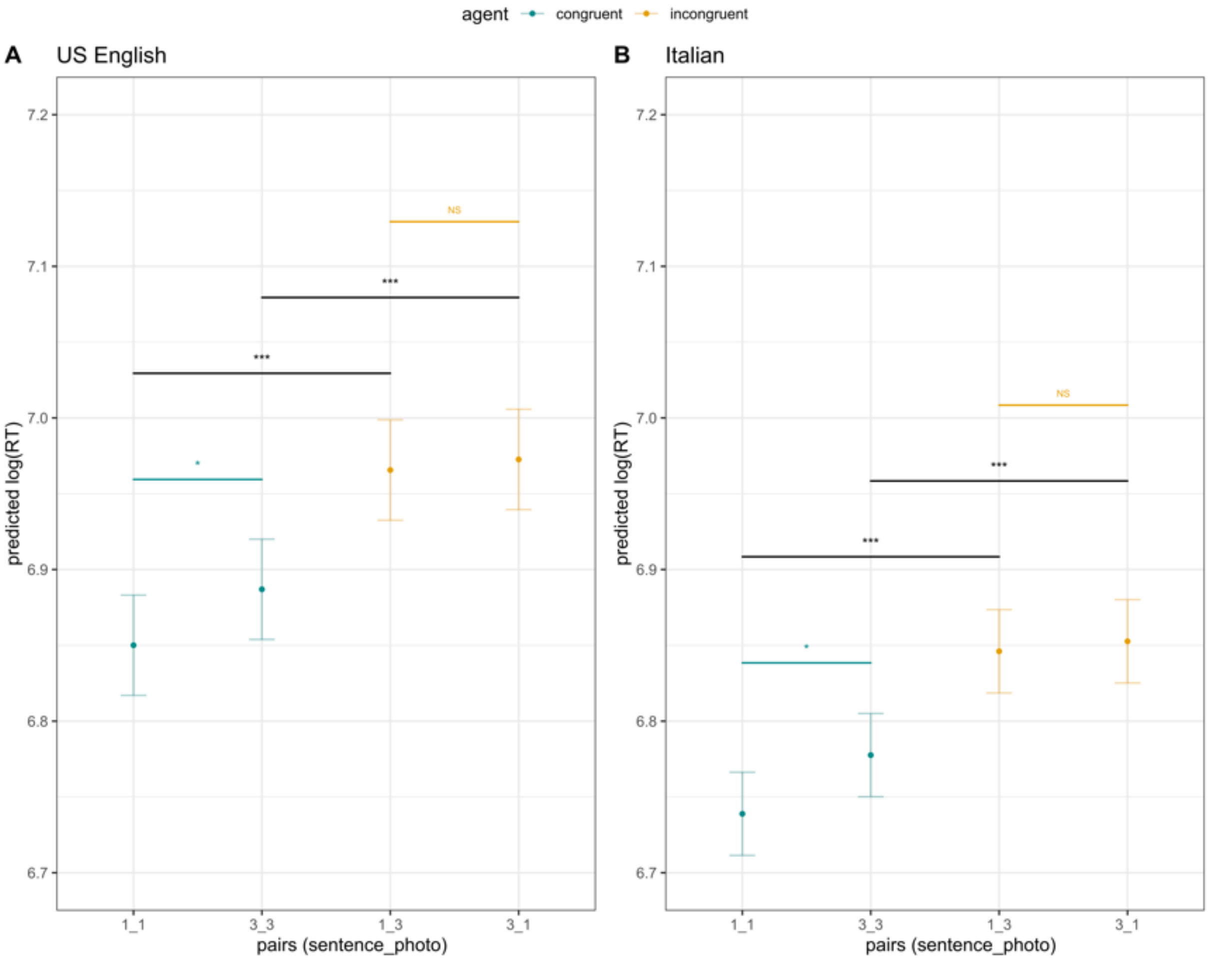

Figure 4 (A) US English sample: cyan color highlights a significant (p = 0.021) difference in favor of 1_1 compared to 3_3 congruent condition; ocher color shows no significant difference between the incongruent sentence-photo pairs; congruent vs. incongruent contrasts are both significantly different (black color; p < .001). (B) Italian sample: cyan color shows a significant (p = 0.022) difference in favor of 1_1 compared to 3_3 congruent condition; ocher color highlights the not significant difference between the incongruent sentence-photo pairs; congruent vs. incongruent contrasts are both significantly different (black color; p < .001). Estimated marginal mean and standard errors are depicted in each graph. 1_1: the reader (i.e., participant) is the agent in both sentence and photo; 3_3: the agent in the sentence and photo is someone else interacting with the reader; 1_3: the reader is the agent of the sentence but occupies the role of receiver of the action in the photo; 3_1: the agent in the sentence is a third person but the agent in the photo is the participant; NS = not significant; * p < .05; *** p < .001

Methods (U.S. English Experiment)

Participants

113 US English speakers' psychology students were recruited for this experiment (M_{age} = 18.79, SD_{age} = 1.84; 20 males, 92 females; 100 right-handed; 2 holding a bachelor's degree, 79 with American high-school degree, and 32 currently enrolled in the bachelor's degree). The sample size used was the same in Italian experiment and online recruitment was carried during the COVID-19

pandemic (fall 2020). Participants were recruited from one of the psychology courses at the University of Hartford and received course credits for their participation in the experiment. Participants gave their consent online to take part in the experiment. The experiment was approved by the Institutional Review Board of the University of Hartford.

Sentence-picture verification task

The same task of the Italian experiment was administered. The sentences were translated by the Italian research team and then checked and corrected by the USA research team (see Appendix D).

Procedure

The experiment was conducted online following the same procedure used in the Italian experiment. Participants were recruited via the SONA system (https://www.sona-systems.com/default.aspx). Once the experiment was published, participants meeting the preselected inclusion/exclusion criteria could access the Gorilla link to start the task.

Results (U.S. English Experiment)

Statistical methods were the same as in the Italian experiment. Data processing was the same as well. 29/113 participants were removed as they did not satisfy the 80% accuracy threshold (N = 84 included in the analyses). No stimuli were removed according to the selected criteria (see the homologous section in Italian experiment).

Random and Fixed Effects

The average total accuracies for participants (N = 84) and stimuli were 90% and 94.49% respectively. Overall, 6142 observations were included (1_1 = 1578; 3_3 = 1551; 1_3 = 1495; 3_1 = 1518) after removing incorrect responses and filtering reaction times (RTs between 100ms and 3000ms).

In the first block of analysis, I put participants and stimuli pairs as a RE and the variables Sentence Agent and Photo Agent as fixed effects. Both RE were found to be significant ($p < .001$) and represented 44% of the variance (participant ICC = 43.4%, stimuli ICC = 0.6%) in the dependent variable (i.e., picture verification). This again indicates that most of the variability in the RT was due to participant differences.

Confirming and replicating the results from the Italian experiment, findings indicated a significant interaction (Photo Agent by Sentence Agent) ($F_{1, 76} = 81.34$, $p < 0.001$, $\eta^2_p = 0.52$, 95%CI[0.36, 0.63]). Main effect of Photo Agent was not significant, although Sentence Agent showed a statistical tendency ($p = 0.053$). Results indicated that, when the participant was the agent in the sentence, image verification times were faster when the participant was the agent in the photo ($M_{log} = 6.85$, $SE_{log} = 0.03$) than in images where the participant was the observer of someone else acting ($M_{log} = 6.97$, $SE_{log} = 0.03$). Conversely, when the agent in the sentence presented before the image was someone else, verification times were faster for images in which the agent was someone else ($M_{log} = 6.89$, $SE_{log} = 0.03$) than ones in which the agent was the participant ($M_{log} = 6.97$, $SE_{log} = 0.03$). See Figure 3 for this result (left side).

In the second block of analysis, I put participants and stimuli pairs as a RE and the variable Sentence-Photo Pairs as a fixed effect (4 levels within-subjects design). Both RE were found to be significant ($p < .001$) and represented 43.98% of the variance (participant ICC = 43%, stimuli ICC = 0.98%) in the dependent variable. Again, RE are in line with the Italian sample.

Again, as in the Italian experiment, findings indicated significant effect of the Sentence-Photo Pairs ($F_{3, 76.02} = 29.08$, $p < 0.001$, $\eta^2_p = 0.53$, 95%CI[0.37, 0.64]). The results were consistent also when adding hand dominance of the participant in the model (main effect and interaction with Sentence-Photo Pairs effect not significant).

As hypothesized, planned contrast showed a suggestive difference between the congruent conditions 1_1 and 3_3 ($t_{74.4} = -2.35$, $p = 0.021$, $d = -0.27$, 95%CI[-0.50, -0.04]). Specifically, in the condition 1_1 verification times were faster ($M_{log(1\_1)} = 6.85$, $SE_{log(1\_1)} = 0.03$) than in the condition 3_3 ($M_{log(3\_3)} = 6.89$, $SE_{log(3\_3)} = 0.03$). Moreover, I found significant differences in the contrasts between congruent and incongruent conditions. Specifically, 1_1 *vs.* 1_3 ($t_{76} = -7.33$, $p < 0.001$, $d = -0.84$, 95%CI[-1.10, -0.58]) and 3_3 *vs.* 3_1 ($t_{76} = -5.43$, $p < 0.001$, $d = -0.62$, 95%CI[-0.87, -0.38]).

In the congruent condition 1_1 images were verified faster ($M_{log(1\_1)} = 6.85$, $SE_{log(1\_1)} = 0.03$) than in the incongruent condition 1_3 ($M_{log(1\_3)} = 6.96$, $SE_{log(1\_3)} = 0.03$); similarly in the congruent condition 3_3 images were verified faster ($M_{log(3\_3)} = 6.89$, $SE_{log(3\_3)} = 0.03$) than in the incongruent condition 3_1 ($M_{log(3\_1)} = 6.97$, $SE_{log(3\_1)} = 0.03$). There were no significative differences between the incongruent condition 1_3 *vs.* 3_1. These results confirmed and replicated those of the Italian experiment. Moreover, since the results have been replicated in two different languages, they suggest a possible cross-linguistic effect. Again, concerning accuracy no differences in d-prime were found within each pair of the contrasts of interest. Figure 4 (left side) shows the effects found in the sample.

Cross-linguistic Bayesian evidence

To demonstrate the validity of the findings at a cross-linguistic level I used Bayesian statistics Bayes factor bound (BFB) computation. Jeffreys's rule of thumb for BFB interpretation was used (Ly et al., 2016). Evidence from the data in favor of H_1 relative to H_0 (BFB), odds in favor of H_1 relative to H_0, and 'post-experimental odds' combined with prior odds of H_1 to H_0 (set 1:1 for both experiments as I do not have prior odds in favor of a specific hypothesis) were computed as suggested (Benjamin & Berger, 2019).

To test the hypothesis of motor simulation in the agent, I used post-hoc contrasts (1_1 *vs.* 3_3) p-values with Bonferroni correction. To test the hypothesis that condition 1_3 and 3_1 requires both motor simulation and spatial perspective-taking in the agent, the relevant contrasts' p-values with Bonferroni correction were used (as p-values are all $< .001$, I used .001 for the computation). Regarding the motor simulation (1_1 *vs.* 3_3) in the agent, results showed substantial evidence ($1.10 <$ natural log of BF < 2.30; Ly et al., 2016) in favor of H_1 for both studies and a probability of 19% and 18% of H_0 being true for Italian and USA experiments respectively. In addition, given flat prior odds, H_1 is given approximately 4 to 1 for both experiments. For the embodiment in the agent through motor simulation and spatial perspective-taking, findings demonstrate very strong evidence ($BFB > 3.4$; Ly et al., 2016) in favor of H_1 for both studies and conditions and a probability of 2%

of H_0 being true for Italian and USA experiments in all conditions. In addition, given flat prior odds, H_1 is given approximately 53 to 1 in both experiments and conditions. Table 1 shows the BFB, the odds (Pr^U $(H_1|p)$) for H_1 to H_0, and the post-experimental odds.

Table 1 Cross-linguistic Bayesian evidence

Effect	**Language**	**Hypotheses**	**p-value**	**log (BFB)**	**Pr^U $(H_1\|p)$**	**Post-experimental odds**
Agent motor embodiment	Italian	H_1:$\mu_{1\_1} \neq \mu_{3\_3}$ H_0:$\mu_{1\_1} = \mu_{3\_3}$	0.022	1.48	0.81	4.39:1
	US English	H_1:$\mu_{1\_1} \neq \mu_{3\_3}$ H_0:$\mu_{1\_1} = \mu_{3\_3}$	0.021	1.50	0.82	4.49:1
Agent motor and spatial embodiment	Italian	H_1:$\mu_{1\_1} \neq \mu_{1\_3}$ H_0:$\mu_{1\_1} = \mu_{1\_3}$	< 0.001	> 3.97	0.98	53.42:1
	US English	H_1:$\mu_{1\_1} \neq \mu_{1\_3}$ H_0:$\mu_{1\_1} = \mu_{1\_3}$	< 0.001	> 3.97	0.98	53.42:1
	Italian	H_1:$\mu_{3\_3} \neq \mu_{3\_1}$ H_0:$\mu_{3\_3} = \mu_{3\_1}$	< 0.001	> 3.97	0.98	53.42:1
	US English	H_1:$\mu_{3\_3} \neq \mu_{3\_1}$ H_0:$\mu_{3\_3} = \mu_{3\_1}$	< 0.001	> 3.97	0.98	53.42:1

Note: 1_1: the reader (i.e., participant) is the agent in both sentence and photo; 3_3: the agent in the sentence and photo is someone else interacting with the reader; 1_3: the reader is the agent of the sentence but occupies the role of receiver of the action in the photo; 3_1: the agent in the sentence is a third person but the agent in the photo is the participant; BFB: Bayes factor bound; Pr^U $(H_1|p)$: odd for H_1 to H_0; BFB between 1.10 and 2.30 represents substantial evidence in favor of H_1; BFB > 3.4 is indicative of very strong evidence.

General Discussion

Through an action sentence-picture verification task, I aimed to investigate how the processes of spatial perspective-taking and motor simulation interact in sentences involving two characters. I showed that the congruent condition, where the participant is the agent in both the sentence and the picture (1_1), is processed faster compared to the other three conditions (3_3, 1_3, and 3_1). The congruent condition where the agent is someone else in both the sentence and the photo (3_3) is processed slower compared to 1_1, but faster than the incongruent conditions (1_3 and 3_1). In addition, the incongruent conditions are processed slower than the congruent pairs and are not different from each other. Lastly, I demonstrated that the findings are cross-linguistic and occur in at least two different languages, showing common embodied processing.

The key contrast for the study was between 1_1 and 3_3. The fact that 1_1 resulted in faster responses than 3_3 seems to support the hypothesis that the motor simulation takes place in the agent. Indeed, in the 1_1 condition, the subject of the sentence is the agent of the action, therefore motor simulation and spatial perspectives overlap. This is in line with other studies on self-consciousness, which highlight the crucial roles played by first-person agency and spatial perspective in our phenomenology and psychology of the self (Blanke, 2012; Eich et al., 2009; Tversky & Hard, 2009; Vogeley & Fink, 2003). In addition, this finding is in line with Brunyé and co-authors (Brunyé et al., 2009), where first-person ('I') sentences led to the adoption of the first-person (egocentric) perspective.

On the other hand, in the 3_3 condition, I suppose that the reader assumes the spatial perspective of the receiver, prompted by the referential 'me' pronoun (see, spatial grounding hypothesis), but at the same time, she runs a motor simulation as if she were the agent (strong theory on embodiment; e.g., Dectey, 2002). When matching the sentence to the picture, a short delay was registered since the picture corresponds to the sentence in the spatial perspective, but the displayed movement is in the opposite direction with respect to the agency. It is possible that the activation of the motor and premotor cortex, which has been identified as proof of motor simulation during sentence processing (Hauk et al., 2004; Tettamanti et al., 2005), slightly interferes with the movement observed in the picture. Such interference effects between real and simulated movements have already been described in language tasks. For example, some studies have found a selective interference between action words and action execution involving the same effector (Dalla Volta et al., 2009; Liepelt et al., 2012; Mirabella et al., 2012; Nazir et al., 2008; Sato et al., 2008). The findings extended spatial grounding hypothesis findings, demonstrating that the reader assumes the spatial perspective of the receiver, consistently with the syntactic role occupied by the self-referential pronoun, but also simulates the action carried out by the agent in the photo.

Following this line of reasoning, slower reaction times in 1_3 conditions compared to the congruent conditions are accounted for by the violation of both action direction (simulation) and spatial perspective represented in the picture when confronted with the sentence. In the sentence, the reader assumes the agent perspective because of the pronoun ‘I’ in the subject position and simulates the action as such; in the picture, the position of the reader is displayed as the receiver, and the movement is depicted in the correspondent direction. The condition 3_1, however, is unexpectedly as slower as the 1_3 condition, even if in this case the match between sentence and picture should imply only one violation (i.e., the perspective), making it more similar to the 3_3 than to the 1_3 condition. One possible explanation is related to the type of task employed. Considering that the sentence-picture verification task is prominently visual in nature, and therefore the spatial perspective violation may have a greater impact than the motor violation. If this is true, I should expect an opposite pattern of results in a motor task (i.e., 3_3 as slower as 1_3). The task demands have been identified as a key factor to account for contrasting results: for example, the emphasis on the action execution or the imagery of the action can explain inconsistencies in perspective-taking (Pecher et al., 2009; Zwaan & Taylor, 2006). Future studies could address specifically this issue by comparing different kinds of tasks directly.

Finally, the findings suggest that embodied simulations of the agent’s action and spatial perspective are not language-dependent, but seem to be shared in at least Italian and English. Previous work has documented a complex pattern of grammatical similarities and differences in language acquisition of Italian and English (Caselli & Bates, 1999). Consistent differences between the two languages have been found in the noun to verb ratio during adult-to-child speech (Tardif & Shatz, 1997) and in pronominal fluency during English-Italian bilingual learning (Serratrice et al., 2011). To my knowledge, no group has assessed embodied linguistic effects within the context of Italian and US

English. The striking similarities between the two samples may be best interpreted as evidence of cross-linguistic, and potentially universal, embodied processes (cf., Sinha & Kristine, 2001). Note that unlike previous searches for universal structures in the language (cf., Chomskian models), the results point to an embodied universality. Although the results are supportive of such a possibility, an obvious limitation is that the study assessed only two language cohorts. Future research could administer the testing paradigm to other languages. Despite encouraging results, the study has limitations, as I did not consider other psychological confounding variables that could come into play (i.e., accounting for low effect size), like executive functions, egocentric/allocentric spatial preferences, or empathy measures (Brunyé et al., 2016; Gardner et al., 2013; Vukovic & Williams, 2015).

To sum up, this is the first study that I are aware of looking at the relationship between motor simulation and spatial perspective-taking in sentences with two characters and an action. I demonstrated that the two mechanisms are probably, if not entirely, independent of one another and that, from the perspective taken when participants are not the agent of the phrase, motor simulation can take place independently. The existence of a self-referential pronoun can lead participants to choose a certain perspective from among the range of choices that the sentence allows, and I further demonstrated that participants can assume different spatial perspectives during comprehension. Lastly, the common results from the Italian and English samples offer evidence of a cross-linguistic and potentially universal embodied language effect. The findings should be validated by neuroimaging investigations that might demonstrate whether the activation of the motor and premotor cortices arises when utilizing sentences with a third-person agent and the self-referential pronoun as the receiver of an action. In addition, to better comprehend the interaction between action comprehension and spatial information, future studies can look into anomalous agent embodiment in neurodegenerative and mental illnesses that affect social and/or spatial cognition domains (Buckner et al., 2008; Kemp et al., 2012; Serino et al., 2014; Tuena, Mancuso, et al., 2021).

Study 3 - Spatial Frames of Reference and Sentence Simulation

Introduction

Space can be coded, stored, and remembered according to two types of frames of reference: egocentric and allocentric (Burgess, 2008). In the former, space is represented using body-to-object relations, whereas in the latter, the space is represented in an abstract format, independently from our body and location, using objects-to-objects relations. In addition, the egocentric frame of reference allows us to depict our point of view and the allocentric frame of reference allows us to depict the point of view of someone else (Beveridge & Pickering, 2013). These representations can be used in a flexible manner by switching from one to another (Ekstrom et al., 2014). Indeed to adopt the point of view/spatial perspective of someone else, which is not based on our egocentric coordinates, a new set of egocentric coordinates (namely egocentric translocation) are created (Vogeley et al., 2004; Vogeley & Fink, 2003). The ability to create a spatial model of a sentence is possible thanks to situated models (Zwaan & Radvansky, 1998). Thanks to these models, the reader can adopt other characters perspectives other than the one prompted by pronoun referring to him/herself (Beveridge & Pickering, 2013).

To date, only one study focused on the impact of egocentric/allocentric frame of reference, and in particular the proclivity, on sentence simulation (Vukovic & Williams, 2015). By using a VR task, Gramann and co-authors (Gramann et al., 2005) showed that individuals have stable frame of reference preferences to solve the spatial orientation task. In their study (Vukovic & Williams, 2015), they found that when listening to 'You' ("*You are opening a bottle*" - the participant is the agent) and 'I' ("*I am opening a bottle*" - the PC speaker is the agent) sentences, participants with an egocentric spatial orientation preference were faster to verify photos representing their hands acting the listened action compared to photos whit someone else's hands doing the same action. No effect was found for allocentric participants. This effect is altered when a spatial context (i.e., a man with the same perspective of the subject – i.e., who gives the shoulders to the participant – talks to another man in front of the participant) is given and the sentence is heard from PC speakers (i.e., the man speaking). Participants with a preference for the egocentric frame are faster to verify the egocentric photo of the hands performing the listened action after ‘I’ compared to ‘You’ sentences. Participant with a preference for the allocentric frame shows the opposite trend. However, results when the roles of the two men are inverted (now the man that is in front of the participant is speaking), no effect of spatial frames of reference proclivity was found. After ‘You’ sentences, participants were faster to verify egocentric compared to third-person perspective photos, whereas the opposite result was found for ‘I’ sentences. The study shows that: 1) frames of reference proclivity affects sentence simulation, especially for egocentric participants; 2) when combining spatial cognition proclivity with a visual context, results showed that sentences are not simulated adopting the viewpoint of the speaker (the man that acts the sentence) but rather through the participants’ individual perspective. This is similar to what I have found in Study 2. In the 3_3 condition, I argued that the reader assumes the spatial perspective of the receiver, prompted by the referential ‘me’ pronoun (see, spatial grounding hypothesis), but at the same time, runs a motor simulation as if the participant were the agent.

Crucially, Ditman and co-authors (Ditman et al., 2010) argued that simulation of concrete action language can occur also during a behavioral task that does not require online processing (like the one used by Vukovic and Williams) of the information (e.g., memory tasks). They found higher recognition accuracy for ‘You’ sentences compared to ‘I’ and third-person action sentences, showing that simulation, through an enactment effect, can occur spontaneously in memory tasks when the participant is the agent.

In the present experiment I sought to explore if individual egocentric/allocentric performance in a spatial memory task (regardless of the individual proclivity) predicts action (concrete and abstract action transfer) language recognition in the context of two-characters sentences and offline (i.e., in memory tasks). To pursue these aims, I used a VR spatial memory task to obtain egocentric and allocentric spatial memory scores of the participants and used this performance to see if it influences action-sentence recognition with two characters (agent and receiver). I expect that the egocentric performance, that is the ability of the participant to adopt a new set of egocentric coordinates (i.e., those of the character interacting with the participant in the sentence) and involves sensorimotor information, would mainly affect third-person concrete action sentences; conversely, the allocentric performance, which is an abstract representation of the space detached from bodily coordinates, would affect abstract sentences regardless of the perspective of the action.

Methods

Participants

We recruited 29 Italian young adults for this study (M_{age} = 24.9, SD_{age} = 3.48; males = 15; right-handed = 26; with master's degree = 10, with bachelor's degree = 5, with high-school degree = 13, 1 with high specialization master). Participants were recruited at psychology courses of the Catholic University of Milan. Inclusion criteria were: speaking Italian as native language; age ≤ 30 years. Exclusion criteria were self-reported history of memory and/or language disorders. With a medium Cohen's *d* of 0.5, a power of 0.8, and 80 target stimuli, the power analysis for a mixed-effects model with a stimuli-within-condition (CNC) design (Westfall et al., 2014) required a minimum of 26 participants. The study was approved by the Ethics Committee of the Catholic University of Milan. Participants gave their written consent to participate.

Virtual reality egocentric and allocentric spatial memory task

We employed a landmark-based navigation task to test participants' egocentric and allocentric spatial memory (Guderian et al., 2015). The virtual environment was a circular arena surrounded by a wall, with an obelisk inside of the arena, and some distal cues (i.e., mountain, clouds, sun). Participants could freely navigate with arrows keys and mouse and had to collect eight objects and memorize their locations in the arena (the diameter of the arena was 50 virtual meters). The eight objects were balanced for living (e.g., cat) and non-living (e.g., bike) categories. In the encoding phase, the objects were randomly presented, one at the time. To collect the item and to see the following one, the participants had to navigate to the exact location of the item. Once over it, the object disappeared, and the participant had to find the next one. Object locations could be remembered using the boundaries of the arena (i.e., wall – allocentric cue), an intra-arena landmark (i.e., obelisk – egocentric cue), and distal cues (i.e., mountain range, fixed clouds). Figure 5 shows the environment at encoding. Each object was presented (and collected) 4 times in random order. During the immediate recall phase, participants were shown each item, one at the time, on the bottom part of the PC screen and had to navigate to the exact location where the item (shown) was previously collected. Once there, they had to press the spacebar if they were happy with the remembered location. In random order, either the wall or the obelisk were removed. This forced the use of allocentric (i.e., obelisk removed) or egocentric (i.e., wall removed) spatial memory recall. Each object was tested four times, two forcing the egocentric and two forcing the allocentric spatial frame, for a total of 32 recall trials (16 trials for each allocentric and egocentric recall condition). The response variable was the distance error of each object trial at recall (distance in virtual meters of the recalled position from the actual location at encoding). The greatest error possible is 49 virtual meters.

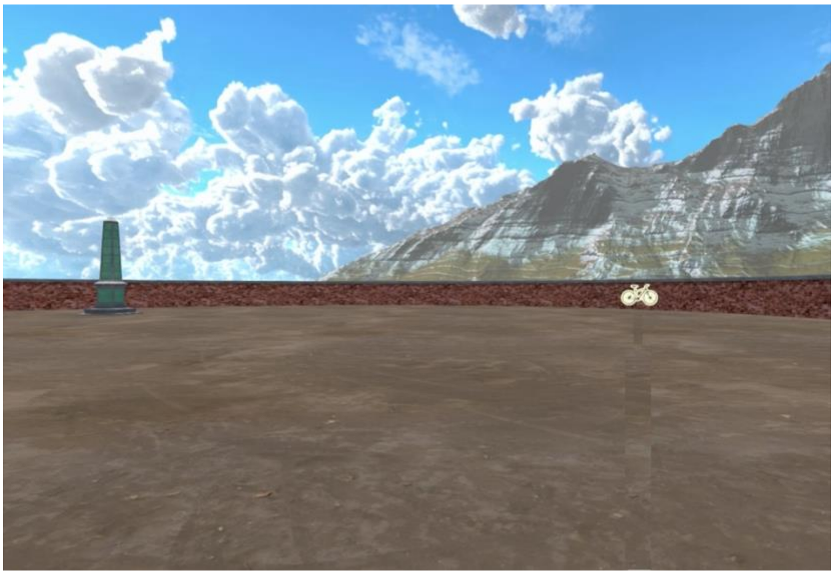

Figure 5 Virtual environment of the arena. Intra-arena landmark (egocentric), wall (allocentric), and distal cues (clouds, mountains) were used to remember the item location.

Sentence recognition memory task

We developed a sentence recognition memory task with Gorilla software (Anwyl-Irvine et al., 2020) with the old-new paradigm (Squire et al., 2007). A set of 160 sentences were created involving two characters (the pronoun 'You' and Gianni or Maria). I used 'You' sentences because this yielded the most consistent results in studies on sentence simulation (Brunyé et al., 2009, 2016; Ditman et al., 2010; Vukovic & Williams, 2015). The set of stimuli was divided into four conditions: action sentences where the participant is the agent, namely first-person perspective action sentences (e.g., "*You shoot the rubber band at Gianni*"); action sentences where the participant is the patient (i.e., the receiver of the action), namely third-person perspective action sentences (e.g., "*Maria shoots you the rubber band*"); abstract sentences where the participant is the abstract agent, namely first-person perspective abstract transfer sentences (e.g., "*You give some time to Maria*"); and lastly, abstract sentences where the participant is the abstract patient, namely third-person perspective abstract transfer sentences (e.g., "*Maria gives you some time*"). Forty action-sentences were taken from Study 2 (Italian experiment), the remaining were created from scratch. The pronouns (in Italian) were always in the first or last place to reduce any influence of the pronoun position.

As in Díez-Álamo and colleagues (Díez-Álamo et al., 2020), I created a main list of 80 sentences (20 sentences for each condition) and then created a reversed list (i.e., the sentence "*You shoot the rubber band at Gianni*" in the main list, was reversed *"Gianni shoots you the rubber band"*). Following the procedure of Díez-Álamo and colleagues, participants in the encoding phase learned the sentences of the main list (old items); whereas in the immediate recognition phase, the sentences of the reversed list were shown as lures/new items in addition to old items. I counterbalanced this aspect across participants, so that the participant that learned "*You shoot the rubber band at Gianni*" had as a new item *"Gianni shoots you the rubber band"*; the following participant had as target *"Gianni shoots you the rubber band"* and as new item "*You shoot the rubber band at Gianni*". Sentences were balanced also by gender (Maria/Gianni) so that half of the action and abstract sentences involved a male and half a female. In addition, sentences length was not statistically different for action and abstract sentences, in the main and reversed list, between first and third-person perspective.

To reduce the effort to learn 80 sentences in only one block and then recognize 160 items, I divided the recognition task into two identical parts. In this way in the first block, participants learned 40 target items and recognized 80 (old plus new) sentences; then a second block was presented, where the remaining 40 target sentences were learned, followed by the recognition part. To reduce any potential order of presentation effect, the blocks were counterbalanced across participants. In this way, old-new (first vs. third-person perspective) items and presentation blocks were counterbalanced during the task yielding four possible lists to be administered. See Appendix E for an example of one of the four lists presented.

In the encoding phase, the sentences were presented at the center of the screen for 5sec followed by a 500ms fixation cross. In the immediate recognition phase, the old-new sentences were presented at the center of the screen and with no time-limit to respond, once responded a 500ms fixation cross appeared. To respond the participants used the mouse by clicking the two relevant (old-new) screen sections placed in the middle of the of the bottom part of the PC screen.

Procedure

The participants were welcomed in a quiet room at the Catholic University of Milan and read and signed the consent form. The PC for this study was a VR-ready Dell G5 15.6inch. Before the tasks, participants were required to provide some demographic information (age, sex, education, handedness). Then the participants were invited to take a seat at a distance of 50cm from the PC screen. In a counterbalanced order the spatial memory and the sentence recognition tasks were administered.

The instructions for the spatial memory task at encoding were: "*Now you will be in a circular virtual arena and your task is to collect some objects and memorize their locations because you will then be asked to remember them later. You will see one object at time. You will see each object four* *times in the same position so that you can remember better its location. To help you to memorize the location you can use the obelisk, the wall, the mountain range, and the fixed clouds as references. You can navigate within the arena with the arrow keys and the mouse. To collect the object, go exactly over it. It will disappear and you will be presented with the next one.*". After this phase, the immediate recognition instructions were read "*You will be asked to put each object in the position where you collected it. However, in random order either the obelisk or the wall will be removed. Once you are in the location you think is correct, press the space bar to release the object and proceed with the next one. You will be asked to replace each object several times regardless of the correctness of your answer*".

The sentence recognition memory task was divided into two blocks (encoding block one and recognition block one, encoding block two and recognition block two). Instructions were displayed before each encoding and immediate recognition phase. Then, the encoding instructions of the first block appeared. "*Now you will see some sentences. Your job is to read the sentences carefully and memorize them because you will then be asked to remember them later. Stay focused because the phrases will only be shown for a few seconds on the screen and will change automatically.*". The instructions of the recognition phase were "*Now you will see some sentences. Your task is to evaluate, using the appropriate buttons on the screen, if the sentence is NEW or OLD. Press OLD with the mouse if you believe that you have seen the sentence among those you memorized a little while ago. Press NEW if you think you have not seen it among those shown before.*".

Statistical analyses

Statistical analyses were carried out with R (version 3.6.3) (R Core Team, 2014). Generalized (logistic) linear mixed-effect model (GLMM) ANOVA with Type III Wald chi-square test was used to explore the effect of the two categorical (action sentence type: concrete or abstract; sentence agent: 'You' or a third person) predictors and the egocentric and allocentric spatial memory error on the correctness (0 = incorrect; 1 = correct) old-new responses. More specifically two separate GLMM were analyzed, one for the egocentric error and the other for the allocentric error. Bobyqa optimizer was used to check model convergence (Brown, 2021). The R formula was as follow [glmer(correct ~ type*agent*egocentric/allocentric error + (1|participant) + (1|sentence), family = binomial)]. Due to convergence issues, I used only a random intercept structure for participants and stimuli. Simple slope analysis with Bonferroni correction (*emmeans* package) was used in the case of a categorical by continuous interaction (Lenth, 2018). I employed phi (φ) as a measure of effect size of chi-square test, which is interpreted as the r coefficient of correlation (Fleiss, 1994). The intraclass correlation coefficient (ICC) was used to examine the impact of the random effects on the dependent variable. Assumptions of logistic GLMM assured by visual inspection (linearity in the logit for continuous variables, absence of multicollinearity, and lack of strongly influential outliers). Data pre-processing was carried out by removing outlier (inter-quartile range method) observations depending on the levels of sentence type and agent (removed 312/4640) for the sentence recognition task and depending on the landmark (removed 36/800) for the spatial memory task. Four participants were removed from the analyses as due to technical problem during the VR task their performance could not be saved (only data from the sentence recognition task). I only used hits and false alarms responses to carry out the two separate logistic GLMM. Indeed, in accordance with the signal detection theory (Green & Swets, 1966), a hit represents the probability that the participant reports the signal present when is actually present and a false alarm is the probability that the subject reports the signal present when is absent. In this way it is possible to discriminate the signal (correct items) from the noise (lures). I used these two responses to identify the probability of a correct response using the predicted probabilities of the logistic GLMM. In this way, I extracted for each response a probability of correctness, depending on the fixed and random effects put in the model, which ranged from 0 to 1. The cut-off of 0.5 was used to define the response as incorrect (≤ 0.5) or correct (> 0.5). Then, I used this categorical predicted response to plot the probability function and the predictive accuracy of the model.

Results

Memory tasks accuracy performances

Overall sentence recognition accuracy for the participants (n° hit/total stimuli for each participant) was 60.7 % (SD = 10.56) and for the sentences (n° hit/total stimuli for each sentence) was 60.7% (SD = 11.1). Given the variability (SD) in accuracy across participants and stimuli, I preferred to use a method that in the computation accounts for such variability and opted for a logistic GLMM (see logistic GLMM section below) with participants and stimuli as random effects.

First of all, I analyzed the accuracy performances in the two memory tasks separately. I investigated the effect of the conditions (Sentence Agent and Sentence Type) with a 2x2 within linear mixed-effects ANOVA (Luke, 2017). Regarding recognition accuracy as measured by d- prime (Green & Swets, 1966), I did not find any main effect of Sentence Agent ($F_{1,72} = 2.62$, $p = 0.11$), of Sentence Type ($F_{1,72} = 0$, $p = 0.99$), or their interaction (F1,72 = 2.25, $p = 0.61$). ICC for the random factor 'participant' was 0.55. This indicates that participants moderately resembled each other and that 55% of the variance in the response variable was due to this grouping structure. Regarding the VR spatial memory task, I used a within (one-way) linear mixed-effects ANOVA (Luke, 2017). I found that a main effect of the Landmark on spatial memory accuracy (virtual error) ($F_1 = 16.54$, $p < 0.001$, $\eta^2_p = 0.02$). The participants were more accurate during the allocentric landmark (i.e., wall) recall condition (M = 10.3, SD = 4.45) compared to the egocentric landmark (i.e., obelisk) recall condition (M = 12.42, SD = 4.53). ICC for the participants was 0.33 and for the objects 0.03. This indicates that the variability in the dependent variable was affected mainly by the random factor 'participant'.

Logistic GLMM models

We ran two separate models, one for the egocentric and one for the allocentric performance.

In the egocentric model, I considered as predictors of correctness (1 = hit, 0 = false alarm) the Sentence Type, Sentence Agent, and the Egocentric Error. I found a significant Sentence Type by Sentence Agent by Egocentric Error interaction ($\chi1 = 4.47$, $p = 0.034$, $\varphi = 0.42$). ICC for the participants was 0.07 and for the stimuli 0.02, showing heterogeneity in the responses across the conditions. Simple slope contrasts showed that the trends of abstract sentences in the first and third-person perspective were not statistically different ($p = 0.571$). Indeed, the probability of a correct response declined, for an increase of one unit in the egocentric error, for the first and third-person perspective abstract sentences respectively by 12.92% ($p < 0.001$) and 11.31% ($p < 0.001$). Conversely, the trends of the concrete sentences in the first and third-person perspective were statistically different ($p = 0.014$). Indeed, the probability of a correct response significantly declined for third-person perspective concrete sentence by 10.42% ($p < 0.001$) for an increase of one unit in the egocentric error but not for the first-person perspective concrete sentence (est. = - 3.61%, $p = 0.355$).

We also found a significant main effect of Sentence Type ($\chi_1 = 10.92$, $p < 0.001$, $\varphi = 0.66$), the Egocentric Error ($\chi_1 = 21.55$, $p < 0.001$, $\varphi = 0.93$), Sentence Agent by Sentence Type interaction ($\chi_1 = 6.33$, $p = 0.012$, $\varphi = 0.5$), and Sentence Type by Egocentric Error interaction ($\chi_1 = 11.14$, $p < 0.001$, $\varphi = 0.68$). The main effect of Sentence Agent ($\chi_1 = 1.04$, $p = 0.309$, $\varphi = 0.2$) and Sentence Agent by Egocentric Error ($\chi_1 = 0,32$, $p = 0.571$, $\varphi = 0.11$) were not significant.

In the allocentric model, I considered as predictors of correctness (1 = hit, 0 = false alarm) the Sentence Type, Sentence Agent, and the Allocentric Error. The Sentence Type by Sentence Agent by Allocentric Error interaction was not significant ($\chi_1 = 3.41$, $p = 0.065$, $\varphi = 0.37$). ICC for the participants was 0.08 and for the stimuli 0.02, showing heterogeneity in the responses across the conditions. Nevertheless, I found a significant Sentence Type by Allocentric Error interaction ($\chi_1 = 7.43$, $p = 0.006$, $\varphi = 0.54$) and a Sentence Agent by Sentence Type interaction ($\chi_1 = 5.21$, $p = 0.022$, $\varphi = 0.46$). Simple slope analyses for the latter did not find any statistically significant differences. However, for Sentence Type by Allocentric Error interaction, I found that the trends of the abstract and concrete action sentences were statistically different ($p = 0.043$). Indeed, the probability of a correct response declined for an increase of one unit in the allocentric error with a greater magnitude for the abstract sentences (est. = -11.77%, $p < 0.001$) than for the concrete sentences (est. = -7.21%, $p = 0.024$).

We also found a significant main effect of Sentence Type ($\chi 1 = 7.5$, $p = 0.006$, $\varphi = 0.55$), Allocentric Error ($\chi 1 = 18.49$, $p < 0.001$, $\varphi = 0.86$), Sentence Agent ($\chi 1 = 3.97$, $p = 0.046$, $\varphi = 0.4$), a Sentence Agent by Sentence Type interaction ($\chi 1 = 5.21$, $p = 0.022$, $\varphi = 0.47$), and a Sentence Type by Allocentric Error interaction ($\chi 1 = 7.43$, $p = 0.006$, $\varphi = 0.55$). The interaction Sentence Agent by Allocentric Error ($\chi 1 = 0,32$, $p = 0.571$, $\varphi = 0.11$) was not significant.

Figure 6 shows the relevant significant results for the egocentric and allocentric models. The use of the cut-off value of 0.5 resulted in an observed vs. fitted model accuracy of 71.02% and 71.14% respectively.

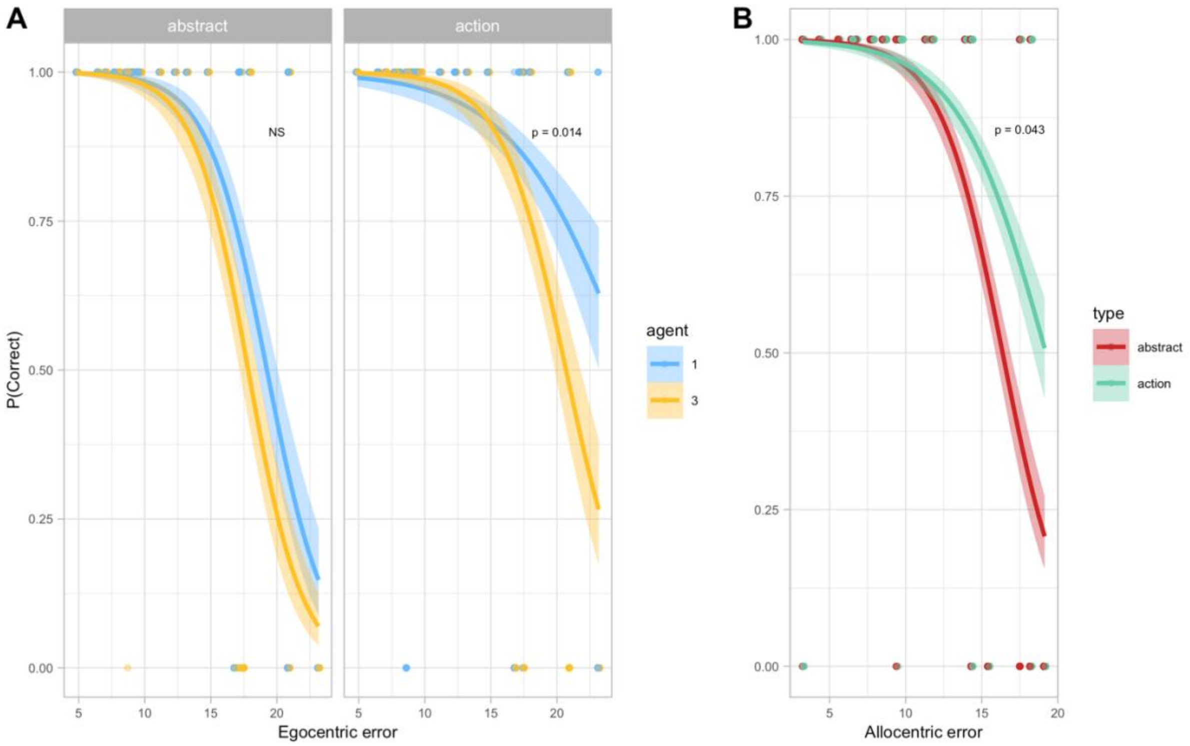

Figure 6 A) significant sentence type by sentence agent by egocentric error contrasts; B) significant sentence type by allocentric error contrast. 1 = first-person perspective; 3 = third-person perspective; P(Correct): predicted probability of correct recognition (hit). NS: not significant. Plot shows the logistic regression lines and relative 95% CI.

General Discussion

In this study, I sought to examine if egocentric and allocentric spatial memory performance had an impact on action language memory, in particular on the recognition of concrete and abstract transfer sentences with two characters, in which the agent is either the participant ('You') or someone else. We found that the egocentric spatial memory accuracy predicts abstract action sentences recognition regardless of the agent, so that the more one fails in the egocentric spatial memory, the more the participant fails in recognizing abstract action sentences; furthermore, the egocentric spatial memory accuracy predicts concrete action sentences recognition when the agent is someone else (i.e., the action is described from the third-person perspective) but not when the agent is the participant (i.e., first-person perspective); the allocentric spatial memory accuracy predicts with a greater extent the correct recognition of abstract than concrete action sentences regardless of the agent. This study demonstrates that there is a link between spatial memory and action language memory with two characters (i.e., agent and receiver) using offline tasks.

Regarding the first finding of this study, I showed that egocentric spatial accuracy predicts the recognition of abstract action sentences regardless of their agent. Despite some authors (e.g., Scorolli et al., 2011) showing that abstract words are acquired and processed through an abstract language system and concrete words by a sensorimotor language system, the finding is more in line with evidence showing that also abstract language is rooted in the bodily experience (Gallese & Lakoff, 2005). Indeed, the ability to remember object locations using the body coordinates predicts the ability to remember abstract actions. It could be possible that some aspects of perceptual and bodily information are used to remember abstract action language and it is predicted by a memory system that uses the body to remember items in the environment.

The second result showed that egocentric memory only predicts the ability to remember concrete action sentences where the agent is someone else. Interestingly, I did not find a significant effect of egocentric memory to remember first-person perspective concrete action sentences, rather I showed that this is true for third-person perspective concrete action language. This finding is not in line with

Vukovic and Williams (Vukovic & Williams, 2015) that showed that individuals with a preference for egocentric spatial processing process faster first compared to third-person perspective concrete action sentences. However, this involved sentences with only one character and concrete actions. It is possible that the result is linked to the ability to adopt a new point of view for a concrete action sentence. Indeed, Vogeley and colleagues (Vogeley et al., 2004; Vogeley & Fink, 2003) showed that to adopt a third-person perspective a new set of egocentric coordinates is required through the egocentric translocation of current own-body to someone else-body coordinates. It could be possible that low ability in egocentric memory predicts this egocentric translocation process.

Then, I showed that allocentric memory predicts abstract action sentences more than concrete action sentences. Allocentric memory is a long-term memory representation of the space (Byrne et al., 2009) and is defined as an abstract cognitive map of the environment (Burgess et al., 2002; O'Keefe & Nadel, 1978). Moreover, Vukovic and Williams (Vukovic & Williams, 2015) showed that individuals with a preference to represent the space with allocentric coordinates do not have a preference to simulate first or third-person concrete action sentences. The finding is in line with such results. Indeed, allocentric memory predicts to a greater extent the ability to recall abstract action sentences and to a lesser extent the recognition of concrete action language regardless of the agent (first vs. third-person).

Lastly, I showed that spatial memory ability is linked to offline processing, and possibly simulation, of action language. Ditman and colleagues (Ditman et al., 2010) showed that participants remember better (d'prime) single character concrete action sentences from their perspective ('You are slicing a tomato') compared to 'I' and 'he/she' pronouns. They argued that this is due to an enactment in memory of the action. Surprisingly, using d'prime, I did not find any significant differences. Here, in the context of two characters' sentences, there were not differences in d'prime depending on the action sentence type and agent. However, the difference in the sentence types was found to be moderated by egocentric and allocentric performance using GLMM. It is possible that the participant, to remember the item, simulates the content of the sentence and particularly, this process could be predicted by egocentric ability for third-person concrete action sentences. Indeed, in accordance with the spatial grounding hypothesis (Beveridge & Pickering, 2013), in absence of spatial cues that help to build the situated model, the reader can simulate motor aspects of text and adopts the perspective of the agent (a third-person agent). Nevertheless, this study has some limitations. First of all, the sample size is smaller than the one required by the sample size computation (25 instead of 26). However, the effect sizes reported with 25 participants reached the effect size required (i.e., medium). In addition, the present study lacks neurophysiological measures that could have improved the understanding of the neural basis of motor and perspective simulation during the recognition task in function of the spatial memory assessment. Future studies could overcome such limitations by employing a larger sample size and/or neurophysiological measures. To conclude, this study shows that spatial memory ability is crucial for simulating action language, in particular, to adopt a third-person concrete action perspective possibly by means of a process of egocentric translocation.

Study 4 - Spatial Distance and Phonological Similarity

Introduction

Two complementary conceptual mappings were proposed by Lakoff and Johnson (1999) to define the link between space and similarity: spatial closeness as similarity and spatial distance as dissimilarity or alternatively, similarity as spatial closeness and dissimilarity as spatial distance. A substantial body of research has demonstrated the significance of space for the representation of similarity, particularly in the semantic domain (i.e., the conceptual metaphor direction explored here is ‘spatial proximity is similarity’) (Casasanto, 2008; Guerra & Knoeferle, 2014). Conversely, Boot and Pecher (2010), presented evidence in support of the opposite direction of the metaphor (i.e., ‘similarity → proximity’- for further details see ‘Spatial distance as linguistic metaphor’ section). Research mainly focused on semantic aspects of language; however, little is known concerning the interplay between distance in space and other language features. In this study, I focused on phonology (i.e., the abstract representation of speech sounds in a certain language). Current research is showing, in contrast to classic theories (Chomsky, 1980), that also phonology is embodied (Berent & Platt, 2022; Fogassi & Ferrari, 2008; Nathan, 2017). Indeed, perceptual aspects of phonology affect memory performance. Studies on the so-called phonological similarity effect revealed that interference in the phonological store between similar phonological memory traces causes the effect to occur (Baddeley, 1986). In this sense, the immediate serial recall of phonologically similar words is poorer compared to phonologically dissimilar words. Conversely, novel findings suggest that phonological similarity improves memory recall independently from the method used to test the retrieval of the items (Gupta et al., 2005). Phonological similarity affects also recognition memory and varies depending on the language (e.g., Chan & Vitevitch, 2009; Luce & Pisoni, 1998; Vitevitch & Rodríguez, 2005). For instance, Spanish words with high frequency of similar words (i.e., neighbourhood) are recognized faster and more precisely than words with low neighbourhood frequency, whereas the opposite result was found in English. Because of this impact on memory, word learning can be influenced by enhancing phonological processing (for example, by focusing on phonological features of words) and pronounceability during encoding, or by phonological similarity (e.g., neighbourhood density) and phonological knowledge (Meade, 2020; Stamer & Vitevitch, 2012).

To my knowledge, no study specifically examined whether the abstract idea of spatial distance is employed offline (stored in memory) to represent words that are similar to or dissimilar from one another in terms of phonology so far.

To achieve this goal, I created a word pairs yes-no recognition task followed by remember-know (RK; Migo et al., 2012; Wixted & Stretch, 2004) and spatial distance judgments. I evaluated the two declarative memory domains, namely episodic memory and semantic memory, using the RK technique (Migo et al., 2012; Wixted & Stretch, 2004).

This study sought to determine if: 1) regardless of the actual encoding spatial location, phonological characteristics of the word pairs will influence subsequent recalled spatial distance; 2) actual and phonological spatial distance estimation of word pairs is affected depending on the declarative memory (episodic memory vs. semantic, R vs. K respectively) system in which this information is stored; 3) phonological characteristics of the word pairs will affect false alarms (FA) in the same way as exposed in the first point (i.e., even when not spatially presented at encoding, similar word pairs are represented closer and dissimilar pairs far apart). To conclude, I predict that phonological properties of words, in addition to physical spatial qualities (i.e., real spatial position), maintain abstract spatial information that can be retained in memory as the phonological spatial distance between dissimilar and similar noun pairs.

Methods

Participants

For this study, 61 healthy young adults were enrolled (M_{age} = 23.25, SD_{age} = 4.04, M_{edu}= 14.87, SD_{edu} = 2.76; females = 25; right-handed = 56). Participants were recruited online via Prolific (https://app.prolific.co/) due to the COVID-19 pandemic restrictions during December 2021. Participants were paid 8.73 € per hour (the experiment lasted 20 minutes approx.; range = 13m55sec-22m53sec). Inclusion criteria were Italian as a native language and normal-to-correct vision. Exclusion criteria were self-reported language-related disorders, literacy difficulties, history of head injury, cognitive deficits, amnesia, long-term/chronic disabilities, and psychiatric medications. With a Cohen's d of 0.33, a power of 0.8, 60 stimuli, and a fully crossed design, the power analysis for a mixed-effects model (Westfall et al., 2014) required a minimum of 59 participants. Cohen's d was extracted from the differences between dissimilar and alliterative words at recall (Gupta et al., 2005). Participants gave their consent to participate before the experiment began. The study was approved by the Ethical Committee of the Catholic University of Milan.

Stimuli

For the experiment, 100 two-syllable nouns (with four or five letters) were chosen. I used two-syllable nouns because in Italian this is the minimum number of syllables needed to obtain a pool of meaningful words large enough for the present task. Part of these words (N = 84) was taken from Montefinese and colleagues (2014) database, whereas the remaining (N = 16) were added by the authors and validated using the same dimensions (emotional valence, familiarity, concreteness, and imaginability). The validation study was administered to 21 adults.

Averaged values from the participants for each word were calculated and any outliers were removed. T-tests were carried out to find differences between the dissimilar and alliterative groups and yielded no differences for any dimension between the two groups. In addition, I evaluated, with the CoLFIS database (Bertinetto et al., 2005), the lexical frequency (average frequency for each word pair). I found no difference (t-test) between alliterative and dissimilar phonology group. Alliterative words (the first two phonemes between two words overlap, Gupta et al., 2005; i.e., 'no-ce'/'no-do'; 'nut'/'knot') were considered similar in phonology to each other, whereas words without the same sound on the first two phonemes were considered dissimilar ('fi-ore/to-po'; 'flower'/'rat'). Words were randomly paired to create the alliterative and dissimilar noun pairs. I used alliterative phonemes given their intermediate effect on memory recall (Gupta et al., 2005). Finally, I assessed semantic similarity by calculating the Wu and Palmer index (1994) with Python WordNet and found no difference (t-test) between alliterative and dissimilar phonology group. In the end, only the phonology could distinguish the couples as similar or distinct.

RK spatial distance task

We created an online yes-no recognition task incorporating RK and spatial distance judgements (Migo et al., 2012; Wixted & Stretch, 2004). Thanks to the Gorilla platform for online behavioral research, the task was created (Anwyl-Irvine et al., 2020). The task was broken down into two phases: encoding and immediate recognition. The latter phase was further divided into three sections: a yes-or-no recognition task, an RK judgment task, and a spatial distance estimation task. Forty noun pairs (80 single words) were presented in the encoding phase; additional 20 noun pairs (40 total single words) were used as new items for the recall phase (see Appendix F). The old items were divided into 20 phonologically dissimilar and 20 alliterative noun pairs, similarly, the new items were further categorized into 10 phonologically dissimilar and 10 alliterative noun pairs. The phonologically different and alliterative noun pairs were further classified into far and close (reciprocal) spatial locations throughout the encoding phase (i.e., 10 near-dissimilar, 10 far-dissimilar, 10 near-alliterative, 10 far-alliterative noun pairs). This resulted in four experimental conditions as in a similar study on similarity (Boot & Pecher, 2010; Guerra & Knoeferle, 2014). The near and far locations were defined in the following way. Randomly for each noun pair, a noun was chosen as ‘fixed’ and the other as ‘moving’ (moving across trials). The fixed noun was placed at the center of the bottom section of the PC screen. The moving noun was placed in line with the fixed noun and on the same vertical axis and in each trial, it moved at different distances from the associated fixed word. By using the fixation cross, the near and far portions were extracted. So that if the moving noun was below the fixation cross, it was considered near to the fixed word, conversely if the moving noun was displayed above the fixation cross, it was considered far from the fixed noun. The moving noun of each noun pair was randomly allocated to the far or near portion and this order was counterbalanced among the participants. For example, if the noun ‘noce’ (fixed on the vertical axis bottom location) was paired with the moving word ‘nodo’ in the near portion, for the next participant the noun ‘noce’ was kept fixed at the same location as described above but the moving noun ‘nodo’ was counterbalanced and moved to the correspondent far slot of the screen. So, each fixed noun was paired with a moving noun that shifted across the spatial locations across trials. The noun pairs were presented one at a time for 5sec, followed by a 1sec fixation cross that exactly divided the near and far portions of the screen.

In the encoding phase, participants were instructed to pay attention and memorize the word pairs (intentional encoding), the implicit task was therefore the encoding of the spatial locations of each noun pair. Pairs were randomly presented two times during this phase to improve encoding, as the objective of this study is to maximize the number of retained items for the recognition part. See Figure 7 for the encoding phase flow.

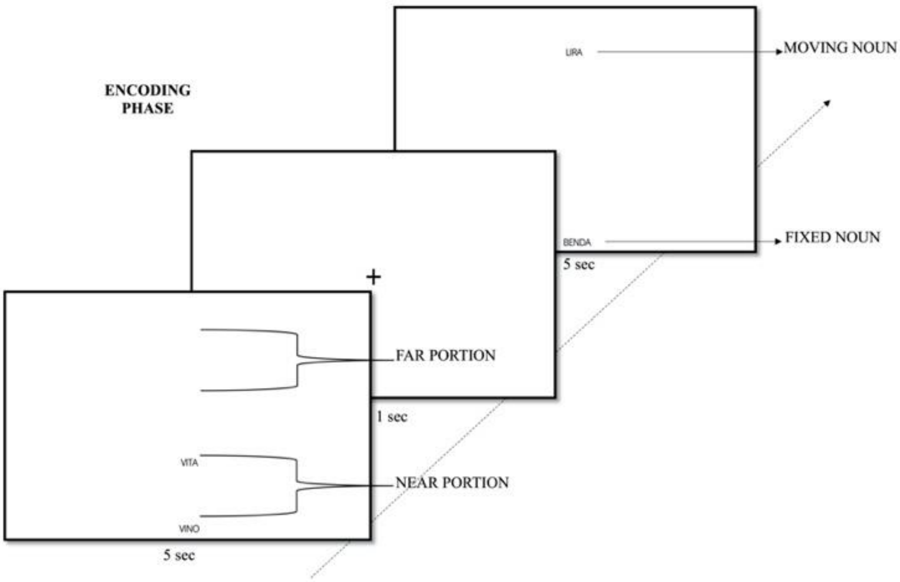

Figure 7 In the encoding phase, the moving noun of each word pair moved between near and far portions across trials.

In the immediate recognition phase, all the old noun pairs and unlearned new item pairs were randomly presented to each participant. The noun pairs were shown on the horizontal axis equidistantly from the fixation cross to avoid any bias regarding the encoded vertical position. Participants were instructed to judge the noun pairs as old or new. If they rated the words as old, they were asked if they ‘remember’ or ‘know’ the noun pairs according to the RK procedure (Migo et al., 2012). For the old-new and RK responses, a 2-forced choice method was applied (no time limit to respond), and participants used the mouse to select the preferred box shown on the PC screen. After both R and K responses, they had to judge the distance of the moving noun from the fixed one. They could move the slider tooltip (i.e., position marker) with the mouse and once happy with the slider tooltip position, they pressed the spacebar to proceed to the following noun pair (no time limit to respond). In particular, the participant was presented with a vertical slider (underlying not visible values 0 to 5 by 0.1) covering the distance between the near and far portions, where the 2.5 slider value corresponded to the fixation cross. The starting position of the slider tooltip was in the middle of the slider range (i.e., the location of the fixation cross). The fixed word appeared at the same location as in the encoding phase. The participant was asked to move the slider tooltip to indicate where the second word was located during the encoding phase. Importantly, the spatial judgment task was unexpected as participants were instructed only to memorize the word pairs and not their spatial location on the PC screen. Lastly, if the participant responded ‘new’, the following noun pair was presented, skipping the RK section of the task (and consequently the spatial judgment task). A 1sec fixation cross was presented after each new response or spatial distance judgment response. See Figure 8 for the immediate recognition phase procedure. In addition, 13 total attentional checks (“*press the box with the number 1*” –button box ‘1’ or button box ‘10’ choices) were put across the encoding and recognition phases according to suggestions provided by Gorilla (Anwyl-Irvine et al., 2020).

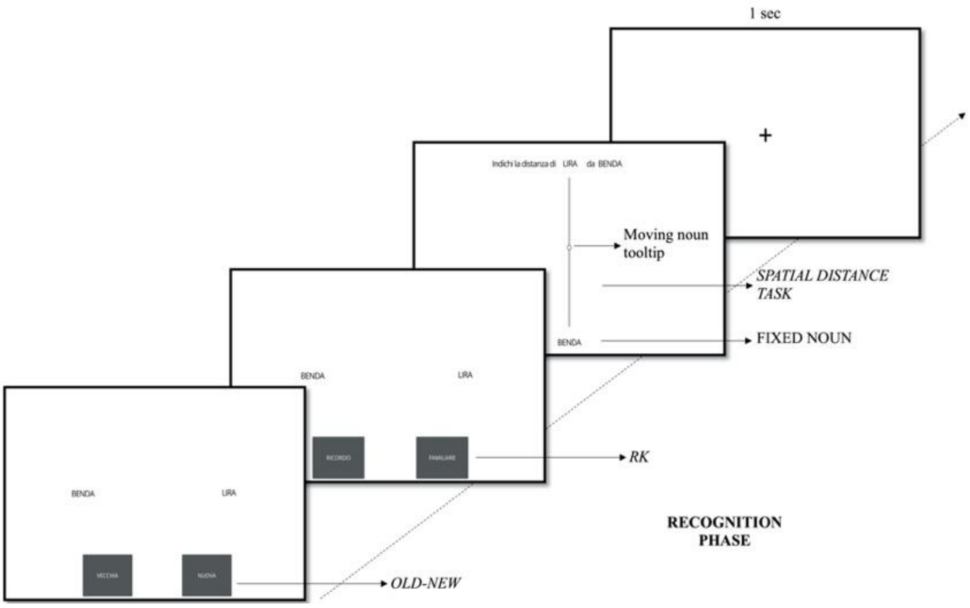

Figure 8 Procedure of the remember-know (RK) spatial distance judgment task. In the recognition phase, old-new (first screen from left), RK (second screen from left), and spatial distance tasks (third screen in the middle) were performed

Procedure

Due to the COVID-19 epidemic, the experiment was carried out online. When the experiment was posted on the Prolific system, individuals who met the predetermined inclusion/exclusion criteria may access the Gorilla link to begin the experiment. After ticking the consent form (mandatory for proceeding further), demographical (i.e., age, sex, education, dominant hand) information was collected. Then, participants read the following instructions: "*Now you will see some word pairs. Your task is to memorize the word pairs. Pay attention because each pair of words will only be visible for a few seconds. You will see each pair of words twice in random order throughout the presentation. The word pairs will be shown automatically*". After the encoding phase, the recognition phase instructions were displayed: "*Now you will see some word pairs and you will have to indicate if: the pair is old, that is, if you saw it among the pairs of words presented before ('OLD' button); the pair is new, that is to say, that you have not seen it among the pairs of words presented before ('NEW' button. If you answer 'OLD', you will be asked if: you remember the pair of words, that is, if you have a detailed memory of the noun pair ('REMEMBER' button); the pair is familiar to you, i.e., if you know you have seen them but do not have a detailed memory of it*

('KNOW' button). Once indicated if you remember/know the word pair, you will also have to indicate the distance of the words using a slider". During this latter phase on the top section of the screen of each spatial judgment, the participant is prompted to indicate the distance of the moving from the fixed noun (the actual nouns are presented, e.g., "*indicate the distance of NODO from NOCE"*). Any questions or technical issues could be resolved thanks to the Prolific chat with the principal investigator of the experiment.

Statistical analyses

Statistical analyses were carried out with R (version 3.6.3) (R Core Team, 2014). Linear mixed-effect model (LMM) ANOVA using Satterthwaite approximation was applied in the analyses (Luke, 2017). Bobyqa optimizer was used to check model convergence (Brown, 2021). Models were specified to have random intercept for noun pairs and participants, as random intercept and slope models failed to converge [formula = outcome ~ fixed effects + covariate + (1|participant) + (1|noun pair)]. The encoding of the noun pairs' phonological similarity (alliterative-dissimilar) and spatial location (near-far) were the fixed effects. To control for different computer screen heights used by the participants that could bias spatial encoding and recalled spatial distance I put height in pixel (px) recorded by Gorilla software as a covariate in each model. This resulted in a 2x2 within subjects ANCOVA. Variance explained by random effects on the dependent variable was provided by the intraclass correlation coefficient (ICC). *lme4* R package was used to run the LMM analyses (Bates et al., 2015). LMM assumptions of normality of residuals and homoscedasticity were verified by visual inspection. Partial eta squared (η^2_p) was interpreted (small = 0.01, medium = 0.06, and large = 0.14) according to Richardson (2011). To test the amount of evidence for the findings, I also used Bayesian statistics. Bayes factor bound (BFB) computation was carried out as suggested to improve p-value interpretation (Benjamin & Berger, 2019). Jeffreys's rule of thumb for BFB interpretation was used (Ly et al., 2016). Evidence from the data in favor of H1 relative to H0 (i.e., BFB), odds in favor of H1 relative to H0, and 'post-experimental odds' combined with prior odds of H1 to H0 (prior odds were set to 1:1) were computed as suggested (Benjamin & Berger, 2019).

Regarding pre-processing, to check that participants were giving above chance responses in the recognition phase (old-new), d-prime (d') was used to find participants with below-chance performance (i.e., d' ≤ 0) (Green & Swets, 1966). Two participants were found to have random guessing performance and were excluded. Concerning responses of the old-new, RK, and spatial judgment tasks exceeding 500ms or 5000ms were excluded and coded as missing values as these can be handled properly by LMM (Brown, 2021). Correct rejection (CR) had no missing values out of 1029 responses, miss had 38 out of 544 responses coded as missing values. Regarding RK FA, 20/123 responses were coded as missing values and regarding the spatial distance judgment task 10/123 were coded as missing. Concerning R hit, 293/1209 responses were coded as missing values and regarding the spatial distance judgment task 133/1209 were coded as missing; for K hit, 75/607 responses were coded as missing values and regarding the spatial distance judgment task 54/607 were coded as missing. All the participants responded correctly to the attentional checks (range 12-13 out of 13). Values in the graphs and result section are the predicted values of the LMM. The significance level for all the analyses was set to 0.05.

Results

Recognition accuracy performance

Mean recognition memory performance (n° hit/40) for the participants and conditions was 77% (SD = 23.13). Table 1 shows the old-new performances by condition in detail.

The accuracy of the spatial distance task was extracted in the following way. The values between 0 and 2.49 represent the near portion, whereas values between 2.51 and 5 the far portion (2.5 is the position of the fixation cross). If the participants put the slider tooltip in the near portion and the moving noun at encoding was in the near to the fixed noun it was coded as correct, if the participants put the slider tooltip in the far portion and the moving noun at encoding was in the far section of the screen relative to the fixed noun position again it was coded as correct. Conversely, if the slider tooltip was placed in the opposite portion of the screen, responses were recorded as incorrect (e.g., if the slider tooltip was placed above the fixation cross – far portion – but at encoding the moving noun was below the cross – near to the fixed noun – the response was incorrect). Table 2 shows the spatial distance performances by condition.

Table 2 Mean recognition memory and spatial distance task accuracy

Old-new task accuracy performance				
Response	*far-alliterative*	*far-dissimilar*	*near-alliterative*	*near-dissimilar*
Hit	7.83 (2.26)	7.42 (2.53)	7.88 (2.11)	7.64 (2.37)
Miss	2.92 (2.13)	3.64 (2.17)	3.05 (1.86)	3.2 (2.1)
Response	*alliterative*		*dissimilar*	
CR	9.31 (1.18)		8.14 (2.14)	
FA	2.85 (1.72)		2.07 (1.27)	
Spatial distance task accuracy performance				
Response	*far-alliterative*	*far-dissimilar*	*near-alliterative*	*near-dissimilar*
R correct spatial recall	156/267 (missing = 30)	175/268 (missing = 37)	217/277 (missing = 27)	189/274 (missing = 29)
K correct spatial recall	70/151 (missing = 14)	72/121 (missing = 12)	0/143 (missing = 18)	0/138 (missing = 10)

Note: in the old-new section mean and SD are reported. CR and FA do not have spatial encoding positions because were not shown in the encoding phase of the task. Missing are responses exceeding 500ms and 5000ms. CR: correct rejection; FA: false alarms; R: Remember; K: know

Concerning the d' on the old-new responses according to the encoding spatial location (near-far) and noun pairs phonology (dissimilar-alliterative), I only found a main effect of phonology (F1 = 39.83, $p < 0.001$, $\eta^2_p = 0.19$, 95%CI[[0.09, 0.29]). Higher recognition accuracy for dissimilar (M = 2.26, SE = 0.09) than alliterative word pairs (M = 1.93, SE = 0.1) was found. No effect of encoding spatial location or interaction effect was found. Importantly, the average d' for the four conditions exceeded the cut-off (d' ≤ 0) of random guessing. The d' for the far-alliterative condition was 1.92 (SD = 1.04), for the far-dissimilar was 2.24 (SD = 0.96), for the near-alliterative was 1.94 (SD = 1.09), and for the near-dissimilar was 2.29 (SD = 0.94). I also found a main effect on phonology on the beta parameter (observer's bias to say 'old' or 'new') ($F1 = 31.46$, $p < 0.001$, $\eta^2_p = 0.19$, 95%CI[[0.1, 0.3]). In particular, higher bias to respond 'old' for dissimilar ($M = 2.19$, $SE = 0.12$) than alliterative word pairs ($M = 1.46$, $SE = 0.1$) was found (a beta close to 1 represents an unbiased response).

Spatial distance task performance

LMM ANCOVAs were used to assess the impact of spatial distance (near-far) and phonology (dissimilar-alliterative) conditions at encoding by controlling for computer screen height (px) on the recalled spatial distance. Separate within LMM ANCOVAs (word phonology: 2 levels; encoding spatial location: 2 levels; screen height as a covariate) were used to analyze R and K responses. Regarding the spatial distance judgments of R responses, ICC for the random effects was 0.07. I both found a significant effect of the encoding spatial location ($F_1 = 121.66$, $p < 0.001$, $\eta^2_p = 0.11$, 95%CI[0.07,0.14]) and a significant main effect of phonology ($F_1 = 8.23$, $p = 0.007$, $\eta^2_p = 0.18$, 95%CI[0.02,0.39]). A significant effect of the covariate screen height was found ($F_1 = 4.97$, $p = 0.031$, $\eta^2_p = 0.11$, 95%CI[0.00,0.31]). Regarding the main effect of encoding spatial distance, noun pairs that were encoded as near were recalled closer ($M = 2$, $SE = 0.01$) compared to noun pairs that were encoded as far ($M = 2.88$, $SE = 0.01$). Regarding the main effect of noun pairs phonology, alliterative noun pairs are recalled closer ($M = 2.28$, $SE = 0.02$) than the dissimilar noun pairs ($M = 2.59$, $SE = 0.02$). Regarding the covariate, the higher screen in pixel the shorter is the recalled spatial distance. No interaction effect between phonology and spatial location was found. See Figure 9 for these main effect findings.

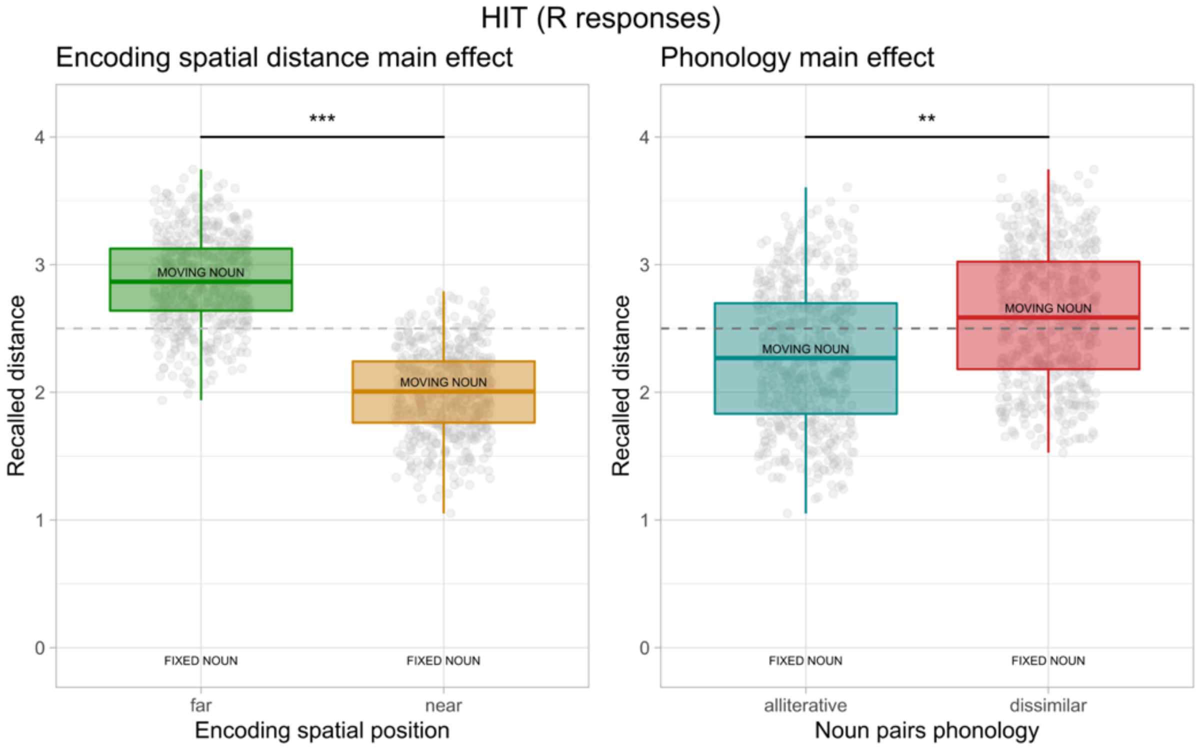

Figure 9 Spatial distance judgments results for the hits after remember (R) responses. The dashed line represents the boundary of the near and far portions of the screen relative to the fixed noun (i.e., fixation cross position). Boxplots depict the recalled distance (range 0-5) of the moving from the fixed word during the spatial distance task. ** < 0.01; *** < 0.001

Concerning the spatial distance judgments of K responses, ICC for the random effects was 0.04. I found only a significant effect of phonology ($F_1 = 13.7$, $p < 0.001$, $\eta^2_p = 0.03$, 95%CI[0.01,0.06]). Alliterative word pairs were recalled as closer (M = 2.33, SE = 0.01) compared to dissimilar noun pairs (M = 2.67, SE = 0.01). Again, no interaction effect between phonology and spatial location was found. See Figure 10 for the results of K responses.

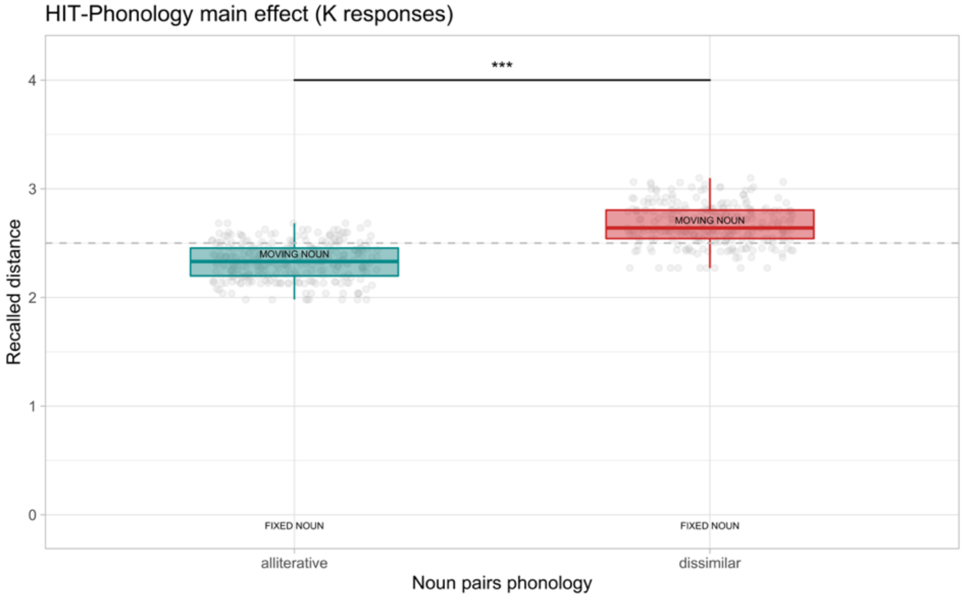

Figure 10 Spatial distance judgments result for hits after know (K) responses. The dashed line represents the boundary of the near and far portions of the screen relative to the fixed noun (i.e., fixation cross position). Boxplots depict the recalled distance (range 0-5) of the moving from the fixed word during the spatial distance task. *** < 0.001

As a second step, FA were considered. Concerning the spatial distance estimation after R, only 18 (six for dissimilar pairs) responses were given by the participants and the observed power of the LMM ANCOVA was 10%. Hence, analyses were not carried out. Regarding the spatial judgments after K, ICC for the random effects was 0.08. I found a significant effect of the phonology of nouns ($F_1 = 9.32$, $p = 0.003$, $\eta^2_p = 0.11$, 95%CI[0.01,0.26]). Alliterative noun pairs were judged to be closer (M = 2.34, SE = 0.03) compared to the dissimilar noun pairs (M = 3.16, SE = 0.05). The height (px) covariate was not significant as they did not see the word pairs in the encoding phase. See Figure 11 for this result.

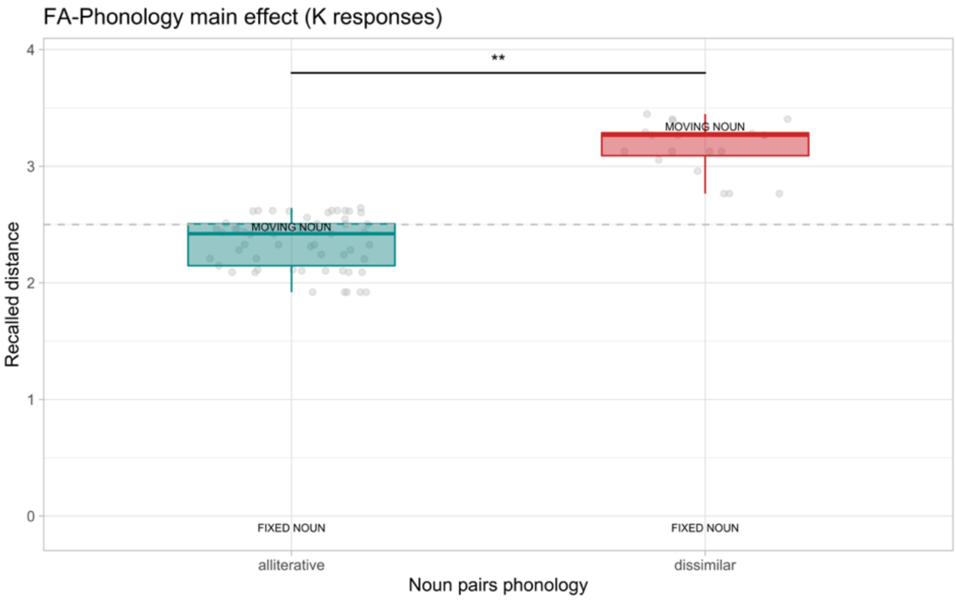

Figure 11 Spatial distance judgments result for false alarms (FA) after the know (K) responses. The dashed line represents the boundary of the near and far portions of the screen relative to the fixed noun (i.e., fixation cross position). Boxplots depict the recalled distance (range 0-5) of the moving from the fixed word during the spatial distance task. ** < 0.01

Bayesian evidence of phonological distance in memory

To strengthen the results, I ran Bayesian computation to ameliorate p-value interpretation (Benjamin & Berger, 2019). These parameters help to understand how much evidence is in favor of the alternative relative to the null hypothesis. Table 3 shows the numeric evidence of the results. Evidence from the data in favor of H1 ($\mu_{\text{phonological dissimilar}} \neq \mu_{\text{alliterative noun pairs}}$) relative to H0 ($\mu_{\text{phonological dissimilar}} = \mu_{\text{alliterative noun pairs}}$) ranges from 21.17 to 186.38 (i.e., strong to extreme evidence; Ly et al., 2016). Odds in favor of H1 relative to H0 range from 0.91 to 0.99 (i.e., the probability of H0 being true ranges from 9% to 1%). Post-experimental odds with prior odds set to 1:1 (H1:H0) were in favor of H1 relative to H0 (see Table 3). Since there is considerable evidence in favor of all the alternative hypotheses (i.e., H1), an abstract phonological spatial distance between words exists and this is particularly evident for K hit responses.

Table 3 *Bayesian evidence in favor of an abstract spatial distance between noun pairs*

Phonological distance	p-value	BFB	Pr^U (H_1|p)	Post-experimental odds
HIT R	0.007	10.62	0.91	10:1
HIT K	< 0.001	> 186.38	0.99	> 186:1
FA K	0.003	21.17	0.95	21:1

Note: Bayes factor bound (BFB) between 10-30 is indicative of strong evidence, BFB > 100 of extreme evidence; Pr^U (H_1|p) is odds in favor of H1 relative to H0. Prior odds for the post-experimental odds were set at 1:1

General Discussion

The goal of this study was to determine whether phonological similarity or dissimilarity between pairs of stimuli is remembered as spatial proximity or distance, namely whether perceptual representation of language includes the conceptual metaphor of space, where dissimilar characteristics are far apart and similar characteristics are close (Lakoff & Johnson, 1999).

We discovered that after hit R responses, spatial position at encoding and phonological similarity of noun pairs has a significant main effect but do not interact. In this regard, I showed that stimuli distance was correctly remembered so that noun pairs that were far at encoding were recalled far regardless of their phonological characteristics and that noun pairs that were near at encoding were judged at recall as near independently of their phonology; parallel to this, I found that the phonologically dissimilar word pairs were recalled as further than the alliterative noun pairs, regardless of their spatial encoding position. The investigations of spatial distance judgements following hits and FA K responses revealed the latter finding. Regarding hit K responses, I did not find a main effect of spatial encoding. Actual spatial information at encoding is not retained for K responses, instead only phonological characteristics drive the spatial distance estimation between the target and moving noun. For the FA K responses, I showed that even when not spatially encoded, phonological features of noun pairs affect spatial judgment estimation, where phonologically similar word pairs are thought to be remembered near to each other, whereas phonologically dissimilar word pairs are thought to be remembered far apart. These three later results (main effect of phonology on distance estimation for hit R, hit K, and FA K) taken together show that, in addition to a remembered physical distance between words (main effect of encoding spatial distance for hit R), exists an abstract spatial distance and it depends on the phonological similarity between the stimuli. Bayesian evidence demonstrates strong evidence of this abstract spatial distance in favor of hit R and FA K responses and extreme evidence in favor of hit K judgments. The effect of phonological similarity of distance judgments is stronger for the latter result, however, I also showed that this effect can influence spatial judgments after R and FA K responses.

Our conclusions on the main effect of phonology on hit R, hit K, and FA K responses' spatial distance estimation extend earlier theoretical and experimental studies. I demonstrated how phonological features of Italian language can facilitate metaphorical conceptualizing (see, Lakoff & Johnson, 1999). Different studies found that proximity can be conceptualized as similarity and distance as dissimilarity (Casasanto, 2008; Guerra & Knoeferle, 2014), however also the opposite direction can be true (Boot & Pecher, 2010). Spatial distance can be used as a metaphor for semantic and perceptual materials (Boot & Pecher, 2010; Casasanto, 2008; Guerra & Knoeferle, 2014; Schneider & Mattes, 2021), I found that space as a metaphor is true also in the context of phonology.

We discovered, intriguingly, that phonological similarity influences hit R and K evaluations. Indeed, the serial parallel independent model (Tulving, 2001; Tulving & Markowitsch, 1998) affirms that perceptual (i.e., the phonological representation of the items), semantic (i.e., whether item pairs are phonologically similar or not), and episodic (i.e., items spatial location) information is encoded serially, stored in parallel, and the retrieval is independent and can entail others systems information. Indeed, R responses has perceptual, semantic, and episodic information and that hit K includes only perceptual and semantic characteristics of the items. Hence, the phonological similarity effect is present in both R and K responses, whereas the spatial information can be accessed only for R responses and independently from other systems. This hypothesis allows also for the independent main effects of phonology and encoding spatial location for hit R judgments.

Concerning FA K, phonological similarity/dissimilarity has driven the spatial judgment of unlearned item pairs. Indeed, typically phonological FA occurs due to surface similarity between words in a list (Chang & Brainerd, 2021), this surface similarity could be used to estimate distances of noun pairs (Boot & Pecher, 2010; Lakoff & Johnson, 1999).

Intriguingly, I also demonstrated that, as indicated by R and K responses, this abstract conceptualization of distance between noun pairs is retained in long-term memory, and especially in episodic and semantic memory. The study by Solomon and colleagues (2019) found that semantic and temporal distance of learned word stored in declarative memory is represented in an abstract cognitive map. I demonstrated that, in addition to semantic information, phonological distance can also be stored in long-term memory with spatial distance information. This might be stored in an allocentric (i.e., object-to-object relations, in this case, the distance between items) *low-dimensional space* (Bottini & Doeller, 2020), where axes are the spatial distance and the phonological similarity of word pairs.

In terms of the main effect of encoding spatial distance, I demonstrated that only hit R but not hit K responses retain the actual physical distance between item pairs. This could be because spatial information is stored in the semantic memory system (Burgess et al., 2002; Eichenbaum et al., 2007) and item-in-context associations are not possible. Liuzzi and colleagues (2019) made the intriguing claim that the perirhinal cortex acts as a connector hub connecting the sensory input of words with more widespread cortical representations of their content. It is possible that the perirhinal cortex, which supports K responses, contributes to this spatial metaphor between phonologically similar and dissimilar words that is based on perception.

Importantly, very low ICC in the LMM showed that the proportion of explained variance in the dependent variable is mainly due to the fixed effects (i.e., spatial distance and phonology fixed effects) (Monsalves et al., 2020). This supports the results regarding the manipulated variables. The Bayesian findings demonstrate strong to extreme evidence in support of the existence of a phonological distance where words are located depending on their phonological features. As a first attempt to demonstrate a relationship between phonology and metaphorical spatial distance, this study does have certain drawbacks. First, the lack of neurophysiological data, like in previous research (e.g., Solomon et al., 2019) on this topic, could strengthen the results. Then, I acknowledge that the manipulation of phonology can be improved and that in Italian orthography is in most of the cases overlapping with phonology (i.e., transparent grapheme-to-phoneme relationships). Additionally, certain participants' performance in terms of recognition memory was low, which resulted in fewer observations for R and K responses and the associated spatial judgments; this is especially true for the FA R condition (see also 95%CI for the effect size in some results).

Here, I demonstrated that, in addition to the semantic and perceptual domains, phonology also bears on the relationship between conceptual similarity and space as a metaphor. According to this research, information and concepts are represented in an abstract cognitive space in addition to a physical space.

Chapter 3: Conclusions

In this work, I outlined some cases of interaction, from an embodied perspective, between two, apparently separated functions, namely language and spatial cognition.

Modularism (Fodor, 1983) posits that the human mind is, at least partially, divided into modules that operate more or less independently of each other. Such a view has been challenged by EC and more recent evidence that shows how cognitive domains are intertwined and sustained by common brain networks (Fincher-Kiefer, 2019; Fogassi & Ferrari, 2008; Spreng et al., 2008; Tettamanti et al., 2005; Vukovic & Shtyrov, 2017). Hence, multiple cognitive domains and brain systems support the formation and retrieval of multimodal and complex cognitive representations (Barsalou, 1999, 2008).

Study 1 (LLE replication) and the unexpected results (LILE instead of LLE in Italian and U.S. English language) provided interesting insights concerning the role of spatial cognition and processing in embodied language. Therefore, three different studies were designed to test and explore the space-language embodied interaction in situated models of action language (Study 2), action language simulation in memory (Study 3), and words phonology (Study 4).

In particular, Study 2 wants to explore 1) the interaction between motor simulation and spatial perspective-taking in action language involving two characters agent and receiver) and 2) if such interaction is cross-linguistically reliable. Study 3 wants to understand how egocentric and allocentric spatial memory performance predicts the recognition of action language involving two actors (agent and receiver). Lastly, study 4 wants to explore if phonologically similar/dissimilar words could be represented metaphorically in memory by spatial distance judgments.

Regarding Study 1, I failed to replicate the LLE. Instead, I found an opposite cross-linguistic (U.S. English, Italian) result (I called this the LILE). Crucially, in a large multicenter study by Morey and colleagues (Morey et al., 2022) it was found that ACE was not replicable in English and non-English speaking participants. Interestingly, they found a consistent main effect of sentence direction, with toward (the body) sentences being processed faster than away (from the body) ones.

Such impressive evidence in favor of LLE is in contrast with the LILE. However, a possible explanation for the results is still possible.

According to previous findings (Beveridge & Pickering, 2013) investigating the reader's point of view while processing a sentence, different pronouns could induce different linguistic perspectives. Following this line of reasoning, I hypothesize that the LILE might be triggered by the spatial perspective assumed by the reader. Such an explanation is not in contrast but complementary to the LLE (Díez-Álamo et al., 2020), as embodiment might depend on several factors, such as individual preferences, task demands, and situation models created by the subject and task instructions. Such a finding raised the possibility that spatial cognition could have an impact in simulating action language when two characters are involved. So, I conducted Study 2 to answer this question.

In Study 2, I found a consistent cross-linguistic (U.S. English, Italian) effect. In particular, I found that when two characters are involved, the reader, during first-person sentences, simulates the action and adopts the egocentric point of view of that action. However, when the agent is someone else, the reader simulates the motor aspect but adopts the spatial perspective of the receiver prompted by the pronoun. I showed that motor simulation and spatial perspective-taking have an additive cost on simulation processes (through RT) during sentence-picture verification tasks. This is in line with the spatial grounding hypothesis (Beveridge & Pickering, 2013) that claims the possibility for the reader to adopt spatial perspectives other than that of the agent (prompted by the relevant pronoun in the text). To further deepen this process, I investigated in Study 3 if the spatial frames of reference ability can influence the memory recognition of action sentences differently depending on the spatial perspective therein described.

In Study 3, I showed that egocentric spatial memory ability (i.e., a body-based representation of the environment), predicts the recognition of action sentences when the agent is someone else (i.e., a third person), whereas allocentric spatial memory ability (i.e., an abstract cognitive map of the environment) predicts abstract sentences regardless of the agent (i.e., participant or someone else). This could hint that it is possible to take the point of view of someone else in a concrete action sentence, but this could be influenced by egocentric spatial ability. Ditman and colleagues (Ditman et al., 2010) showed that during concrete action sentence recognition it is possible an enactment effect that leads to better memory. It could be possible that the ability to remember (and enact) third-person perspectives concrete action sentences is predicted by egocentric ability. Indeed, the process to adopt a new set of egocentric coordinates (a new point of view) requires the so-called egocentric translocation (Vogeley et al., 2004; Vogeley & Fink, 2003). This could be possible also when the task is conducted offline and does not prompt directly simulative processes. This extends the spatial grounding hypothesis (Beveridge & Pickering, 2013), which states that such spatial perspective-taking can occur only when the reader can extract a spatial situated model of the sentence. Indeed, the simulation of sentence recognition has been documented for single-character sentences of concrete actions.

Study 2 and 3 therefore suggest that spatial cognition is a critical aspect that influence situated models and simulation of action language. Previous studies showed that spatial perspective (Brunyé et al., 2009, 2016) is matched with the agent simulation in single-agent sentences and that egocentric/allocentric preferences (Vukovic & Shtyrov, 2017; Vukovic & Williams, 2015) trigger different single character sentence action simulation processes. Here, we showed that the reader while running the motor simulation of the agent could adopt the spatial perspective of the receiver (his/her own) (condition 3_3, Study 2) in two-character sentences. This occurs in a similar manner in at least two languages. In Study 3 we showed that the ability to simulate (offline) the point of view of a third-person agent in two-character sentences is predicted by the reader's egocentric spatial memory ability. The studies demonstrate that spatial cognition is a critical aspect to consider in complex sentences involving two characters and extend previous knowledge on the topic of situated models and action language simulation (Beveridge & Pickering, 2013; Díez-Álamo et al., 2020; Glenberg & Gallese, 2012; Morey et al., 2022; Zwaan & Radvansky, 1998).

A growing number of studies, reported in this thesis, showed that spatial processing is linked to sentence simulation. Little is known if words can have spatial properties depending on their phonological structure. In Study 4, I investigated if phonological aspects of words retain (offline) some physical information (i.e., spatial) that is acquired through the sensorimotor system. Most of the research focused on the interaction between semantics aspects and spatial cognition. In Study 4, I explored if also phonological features of words are linked to a (metaphorical) spatial representation. The results show that this is possible and that single words associations carry spatial-perceptual characteristic that is stored in memory. This extends previous findings on how similarity between concepts is represented using space as a metaphor (Lakoff & Johnson, 1980) and provides further evidence that also phonology and not only meaning is embodied (Berent & Platt, 2022).

This work has some limitations that need to be considered. First, it used only behavioral measures. Neurophysiological data could improve the understanding and the neural underpinnings of simulative (motor and spatial) mechanisms in this context. In addition, due to COVID-19 pandemic restrictions Study 2 and 4 were carried out on-line, in-person testing should be carried out to ensure replicability of the findings. Then, the structure of sentences (from Study 1 to 3) varies across the three studies, and this does not allow a straightforward comparison of the findings. Future studies could better understand how pronouns and sentence structure could influence the effects I have found. In addition, the lack of an U.S. English sample in Study 3 could raise the question of whether the findings are also consistent in a different language, making this experiment less generalizable. It would be interesting to duplicate the method in a sample of American young adults. Lastly, Study 4 has only analyzed the concept of spatial distance as a metaphor, this is a rather rough operationalization of the concept of space compared to the notion of space involved in the Study 2 and 3. Future studies could see if phonology can alter the (egocentric/allocentric) representation of the environment by adopting virtual navigation tasks that enable to assess phonology and spatial cognition.

To conclude, I have demonstrated a series of space-language interaction cases within the framework of EC. The four studies included in this thesis demonstrate that not only space and language are embodied but that possible common embodied representations support both domains and that determine the effects I have shown. The interconnection among cognitive domains, rooted in the sensorimotor experience, requires a more complex approach to the study of single domains, which are interconnected to each other. A multifaceted approach (and methodology) to the study of cognitive psychology and its processes would enable us to grasp the complexity of the human mind.

Acknowledgments

I would like to express my deepest appreciation to my supervisor Prof. Claudia Repetto for her mentoring and teaching. I could not have undertaken this journey without my colleagues Dr. Silvia Serino and Dr. Daniele Di Lernia, who supported me during this period. I would like to express my deepest gratitude to Prof. Giuseppe Riva for his guidance since my master's dissertation. I am also grateful to the referees, Prof. Claudia Gianelli and Prof. Matthew Costello, for reviewing my thesis. Last but not least, a special thanks to my family, girlfriend, and friends for the good times during the arduous ones.

References

Andres, M., Finocchiaro, C., Buiatti, M., & Piazza, M. (2015). Contribution of motor representations to action verb processing. *Cognition*, *134*, 174–184. https://doi.org/10.1016/j.cognition.2014.10.004

Anwyl-Irvine, A., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. (2020). Gorilla in our MIDST: An online behavioral experiment builder. *Behavior Research Methods*, *52*, 388–407. https://doi.org/10.1101/438242

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). https://doi.org/10.1371/journal.pone.0011751

Baddeley, A. D. (1986). *Working memory*. Oxford University Press.

Barry, C., & Burgess, N. (2014). Neural Mechanisms of Self-Location. *Current Biology*, *24*(8), R330–R339. https://doi.org/10.1016/j.cub.2014.02.049

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(4), 577–609; discussion 610-60. https://doi.org/10.1017/S0140525X99252144

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617–645. https://doi.org/10.1146/annurev.psych.59.103006.093639

Bates, D. M., Mächler, M., Bolker, B. M., & Walker, S. (2015). Package Lme4: Linear Mixed-Effects Models Using Eigen and S4. *Journal of Statistical Software*, *67*(1), 1–48.

Benjamin, D. J., & Berger, J. O. (2019). Three Recommendations for Improving the Use of p-Values. *American Statistician*, *73*(sup1), 186–191. https://doi.org/10.1080/00031305.2018.1543135

Berent, I., & Platt, M. (2022). Is Phonology Embodied? Evidence from Mechanical Stimulation. *Journal of Psycholinguistic Research*, *0123456789*. https://doi.org/10.1007/s10936-022-09871-x

Bergen, B., & Wheeler, K. (2010). Grammatical aspect and mental simulation. *Brain and*

Language, *112*(3), 150–158. https://doi.org/10.1016/j.bandl.2009.07.002 S0093-934X(09)00080-7 [pii]

Bertinetto, P. M., Burani, C., Laudanna, A., Marconi, L., Ratti, D., Rolando, C., & Thornton, A. M. (2005). *Corpus e Lessico di Frequenza dell'Italiano Scritto(CoLFIS)*. https://www.istc.cnr.it/en/grouppage/colfiseng

Beveridge, M. E. L., & Pickering, M. J. (2013). Perspective taking in language: Integrating the spatial and action domains. *Frontiers in Human Neuroscience*, *7*(577). https://doi.org/10.3389/fnhum.2013.00577

Birba, A., García-Cordero, I., Kozono, G., Legaz, A., Ibáñez, A., Sedeño, L., & García, A. M. (2017). Losing ground: Frontostriatal atrophy disrupts language embodiment in Parkinson's and Huntington's disease. *Neuroscience and Biobehavioral Reviews*, *80*(January), 673–687. https://doi.org/10.1016/j.neubiorev.2017.07.011

Blanke, O. (2012). Multisensory brain mechanisms of bodily self-consciousness. *Nature Reviews Neuroscience*, *13*(8), 556–571. https://doi.org/10.1038/nrn3292

Bohil, C. J., Alicea, B., & Biocca, F. A. (2011). Virtual reality in neuroscience research and therapy. *Nature Reviews. Neuroscience*, *12*(12), 752–762. https://doi.org/10.1038/nrn3122

Boot, I., & Pecher, D. (2010). Similarity is closeness: Metaphorical mapping in a conceptual task. *Quarterly Journal of Experimental Psychology*, *63*, 942–954.

Borghi, A. M., & Barsalou, L. (2021). Perspective in the conceptualization of categories. *Psychological Research*, *85*(2), 697–719. https://doi.org/10.1007/s00426-019-01269-0

Borreggine, K. L., & Kaschak, M. P. (2006). The action-sentence compatibility effect: It's all in the timing. *Cognitive Science*, *30*(6), 1097–1112. https://doi.org/10.1207/s15516709cog0000_91

Bottini, R., & Doeller, C. F. (2020). Knowledge Across Reference Frames: Cognitive Maps and Image Spaces. *Trends in Cognitive Sciences*, *24*(8), 606–619. https://doi.org/10.1016/j.tics.2020.05.008

Brown, V. A. (2021). An Introduction to Linear Mixed-Effects Modeling in R. *Advances in*

Methods and Practices in Psychological Science, *4*(1). https://doi.org/10.1177/2515245920960351

Brunyé, T. T., Ditman, T., Giles, G. E., Holmes, A., & Taylor, H. A. (2016). Mentally Simulating Narrative Perspective Is Not Universal or Necessary for Language Comprehension. *Ournal of Experimental Psychology: Learning, Memory, and Cognition*, *42*, 1592–1605.

Brunyé, T. T., Ditman, T., Mahoney, C. R., Augustyn, J. S., & Taylor, H. A. (2009). When you and I share perspectives: Pronouns modulate perspective taking during narrative comprehension. *Psychological Science*, *20*(1), 27–32. https://doi.org/10.1111/j.1467-9280.2008.02249.x

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: Anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, *1124*, 1–38. https://doi.org/10.1196/annals.1440.011

Burgess, N. (2006). Spatial memory : how egocentric and allocentric combine. *Trends in Cognitive Sciences*, *10*(12), 553–557. https://doi.org/10.1016/j.tics.2006.10.005

Burgess, N. (2008). Spatial cognition and the brain. *Annals of the New York Academy of Sciences*, *1124*, 77–97. https://doi.org/10.1196/annals.1440.002

Burgess, N., Maguire, E. A., Spiers, H. J., & O'Keefe, J. (2001). A temporoparietal and prefrontal network for retrieving the spatial context of lifelike events. *NeuroImage*, *14*(2), 439–453. https://doi.org/10.1006/nimg.2001.0806

Burgess, N., Maguire, E. a, & O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron*, *35*(4), 625–641. https://doi.org/10.1016/S0896-6273(02)00830-9

Byrne, P., Becker, S., & Burgess, N. (2009). Remembering the past and imagining the future: a neural model of spatial memory and imagery. *Psychological Review*, *114*(2), 340–375. https://doi.org/10.1037/0033-295X.114.2.340.Remembering

Cardona, J. F., Kargieman, L., Sinay, V., Gershanik, O., Gelormini, C., Amoruso, L., Roca, M., Pineda, D., Trujillo, N., Michon, M., García, A. M., Szenkman, D., Bekinschtein, T., Manes, F., & Ibáñez, A. (2014). How embodied is action language? Neurological evidence from motor diseases. *Cognition*, *131*(2), 311–322. https://doi.org/10.1016/j.cognition.2014.02.001

Casasanto, D. (2008). Similarity and proximity: When does close in space mean close in mind? *Memory and Cognition*, *36*(6), 1047–1056. https://doi.org/10.3758/MC.36.6.1047

Caselli, C., & Bates, E. (1999). A comparison of the transition from first words to grammar in English and Italian*. *Journal of Child Language*, *26*, 69–111.

Chan, K. Y., & Vitevitch, M. S. (2009). The Influence of the Phonological Neighborhood Clustering Coefficient on Spoken Word Recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(6), 1934–1949. https://doi.org/10.1037/a0016902

Chang, M., & Brainerd, C. J. (2021). Semantic and phonological false memory: A review of theory and data. *Journal of Memory and Language*, *119*(May), 104210. https://doi.org/10.1016/j.jml.2020.104210

Chemero, A. (2011). *Radical embodied cognitive science*. MIT press.

Chomsky, N. (1980). Rules and representations. *Behavioral and Brain Sciences*, *3*(1), 1–15.

Chrastil, E. R., & Warren, W. H. (2012). Active and passive contributions to spatial learning. *Psychonomic Bulletin & Review*, *19*(1), 1–23. https://doi.org/10.3758/s13423-011-0182-x

Cogné, M., Auriacombe, S., Vasa, L., Tison, F., Klinger, E., Sauzéon, H., Joseph, P., & N'Kaoua, B. (2018). Are Visual Cues Helpful for Virtual Spatial Navigation and Spatial Memory in Patients With Mild Cognitive Impairment or Alzheimer's Disease? *Neuropsychology*, *32*(4), 385–400.

Dalla Volta, R., Gianelli, C., Campione, G. C., & Gentilucci, M. (2009). Action word understanding and overt motor behavior. *Experimental Brain Research*, *196*, 403–412.

De Scalzi, M., Rusted, J., & Oakhill, J. (2015). Embodiment Effects and Language Comprehension in Alzheimer's Disease. *Cognitive Science*, *39*(5), 890–917. https://doi.org/10.1111/cogs.12187

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*, 20–30. https://doi.org/10.1007/s00426-011-0356-1

Dectey, J. (2002). Neurophysiological evidence for simulation of action. In J. Dokicand & J. Proust (Eds.), *Simulation and Knowledge of Action* (pp. 53–72). John Benjamins.

Diefenbach, C., Rieger, M., Massen, C., & Prinz, W. (2013). Action-sentence compatibility: The role of action effects and timing. *Frontiers in Psychology*, *4*(MAY), 1–13. https://doi.org/10.3389/fpsyg.2013.00272

Díez-Álamo, A. M., Glenberg, A. M., Díez, E., Alonso, M. A., & Fernandez, A. (2020). The linguistic looming effect. *Journal of Memory and Language*, *114*(June), 104147. https://doi.org/10.1016/j.jml.2020.104147

Dijkstra, K., & Post, L. (2015). Mechanisms of embodiment. *Frontiers in Psychology, 6:1525*. https://doi.org/10.3389/fpsyg.2015.01525

Ditman, T., Brunyé, T. T., Mahoney, C. R., & Taylor, H. A. (2010). Simulating an enactment effect: Pronouns guide action simulation during narrative comprehension. *Cognition*, *115*(1), 172–178. https://doi.org/10.1016/j.cognition.2009.10.014

Doeller, C. F., King, J. A., & Burgess, N. (2008). Parallel striatal and hippocampal systems for landmarks and boundaries in spatial memory. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(15), 5915–5920.

Eich, E., Nelson, A. L., Leghari, M. A., & Handy, T. C. (2009). Neural systems mediating field and observer memories. *Neuropsychologia*, *47*(11), 2239–2251. https://doi.org/10.1016/j.neuropsychologia.2009.02.019

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience*, *30*, 123–152. https://doi.org/10.1146/annurev.neuro.30.051606.094328

Ekstrom, A. D., Arnold, A. E. G. F., & Iaria, G. (2014). A critical review of the allocentric spatial representation and its neural underpinnings: Toward a network-based perspective. *Frontiers in Human Neuroscience*, *8*(803). https://doi.org/10.3389/fnhum.2014.00803

Engelkamp, J. (1998). *Memory for Actions*. Psychology Press/Taylor & Francis (UK).

Engelkamp, J., Seiler, K. H., & Zimmer, H. D. (2004). Memory for actions: item and relational information in categorized lists. *Psychological Research*, *69*(1–2), 1–10. https://doi.org/10.1007/s00426-003-0160-7

Fincher-Kiefer, R. (2019). *How the Body Shapes Knowledge: Empirical Support for Embodied Cognition*. American Psychological Association.

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

Fodor, J. A. (2000). *The mind doesn't work that way*. MIT press.

Fogassi, L., & Ferrari, P. F. (2008). Mirror neurons and the evolution of communication and language. *Current Directions in Psychological Science*, *16*(3), 133–153. https://doi.org/10.4324/9780203837894

Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, *22*(3–4), 455–479. https://doi.org/10.1080/02643290442000310

Gardner, M. R., Brazier, M., Edmonds, C. J., & Gronholm, P. C. (2013). Strategy modulates spatial perspective-taking: Evidence for dissociable disembodied and embodied routes. *Frontiers in Human Neuroscience*, *7*(JUL), 1–8. https://doi.org/10.3389/fnhum.2013.00457

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). https://doi.org/10.1371/journal.pone.0025036

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception.* Houghton Mifflin.

Giraud, A. L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Neuroscience*, *15*(4), 511–517.

Glenberg, A. M. (2010). Embodiment as a unifying perspective for psychology. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(4), 586–596. https://doi.org/10.1002/ejsp.670

Glenberg, A. M. (2015). Few Believe the World Is Flat: How Embodiment Is Changing the

Scientific Understanding of Cognition. *Canadian Journal of Experimental Psychology*, *69*(2), 165–171. https://doi.org/10.1037/cep0000056

Glenberg, A. M., & Gallese, V. (2012). Action-based language: A theory of language acquisition, comprehension, and production. *Cortex*, *48*(7), 905–922. https://doi.org/10.1016/j.cortex.2011.04.010

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin and Review*, *9*(3), 558–565. https://doi.org/10.1109/TAMD.2011.2140890

Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language*, *26*(1), 69–83. https://doi.org/10.1016/0749-596X(87)90063-5

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*, *61*(6), 905–919. https://doi.org/10.1080/17470210701625550

Gramann, K. (2013). Embodiment of Spatial Reference Frames and Individual Differences in Reference Frame Proclivity. *Spatial Cognition and Computation*, *13*(1), 1–25. https://doi.org/10.1080/13875868.2011.589038

Gramann, K., Müller, H. J., Eick, E. M., & Schönebeck, B. (2005). Evidence of separable spatial representations in a virtual navigation task. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(6), 1199–1223. https://doi.org/10.1037/0096-1523.31.6.1199

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Wiley.

Guderian, S., Dzieciol, X. A. M., Gadian, D. G., Jentschke, X. S., Doeller, C. F., Burgess, N., Mishkin, X. M., & Vargha-Khadem, F. (2015). Hippocampal Volume Reduction in Humans Predicts Impaired Allocentric Spatial Memory in Virtual-Reality Navigation. *The Journal of Neuroscience*, *35*(42), 14123–14131. https://doi.org/10.1523/JNEUROSCI.0801-15.2015

Guerra, E., & Knoeferle, P. (2014). Spatial distance effects on incremental semantic interpretation of abstract sentences: Evidence from eye tracking. *Cognition*, *133*(3), 535–552.

https://doi.org/10.1016/j.cognition.2014.07.007

Gunraj, D. N., Drumm-Hewitt, A. M., & Klin, C. M. (2014). Embodiment during reading: Simulating a story character's linguistic actions. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *40*(2), 364–375. https://doi.org/10.1037/a0034853

Gupta, P., Lipinski, J., & Aktunc, E. (2005). Reexamining the phonological similarity effect in immediate serial recall: The roles of type of similarity, category cuing, and item recall. *Memory and Cognition*, *33*(6), 1001–1016. https://doi.org/10.3758/BF03193208

Hartung, F., Burke, M., Hagoort, P., & Willems, R. M. (2016). Taking perspective: Personal pronouns affect experiential aspects of literary reading. *PLoS ONE*, *11*(5), 1–18. https://doi.org/10.1371/journal.pone.0154732

Hauk, O., Johnsrude, I., & Pulvermuller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, *41*(2), 301–307. https://doi.org/S0896627303008389 [pii]

Huffman, D. J., & Ekstrom, A. D. (2019). A Modality-Independent Network Underlies the Retrieval of Large-Scale Spatial Environments in the Human Brain. *Neuron*, *104*(3), 611-622.e7. https://doi.org/10.1016/j.neuron.2019.08.012

Huffman, D. J., & Ekstrom, A. D. (2021). An Important Step toward Understanding the Role of Body-based Cues on Human Spatial Memory for Large-Scale Environments. *Journal of Cognitive Neuroscience*, *33*(2), 167–179. https://doi.org/10.1162/jocn_a_01653

James, W., Burkhardt, F., Bowers, F., & Skrupskelis, I. K. (1890). *The principles of psychology*. Macmillan.

Jostmann, N. B., Lakens, D., & Schubert, T. W. (2009). Weight as an embodiment of importance. *Psychological Science*, *20*(9), 1169–1174. https://doi.org/10.1111/j.1467-9280.2009.02426.x

Julian, J. B., Keinath, A. T., Marchette, S. A., & Epstein, R. A. (2018). The Neurocognitive Basis of Spatial Reorientation. *Current Biology*, *28*(17), R1059–R1073. https://doi.org/10.1016/j.cub.2018.04.057

Kaschak, M. P., & Borreggine, K. L. (2008). Temporal dynamics of the action-sentence compatibility effect. *Quarterly Journal of Experimental Psychology*, *61*(6), 883–895. https://doi.org/10.1080/17470210701623852

Kemp, J., Després, O., Sellal, F., & Dufour, A. (2012). Theory of Mind in normal ageing and neurodegenerative pathologies. *Ageing Research Reviews*, *11*(2), 199–219. https://doi.org/10.1016/j.arr.2011.12.001

König, S. U., Goeke, C., Meilinger, T., & König, P. (2019). Are allocentric spatial reference frames compatible with theories of Enactivism? *Psychological Research*, *83*(3), 498–513. https://doi.org/10.1007/s00426-017-0899-x

Körner, A., Topolinski, S., & Strack, F. (2016). Routes to embodiment. *Frontiers in Psychology*, *6*(940). https://doi.org/10.3389/fpsyg.2015.00940

Kreiner, H., & Eviatar, Z. (2014). The missing link in the embodiment of syntax: Prosody. *Brain and Language*, *137*, 91–102. https://doi.org/10.1016/j.bandl.2014.08.004

Lakoff, G., & Johnson, M. (1980). The metaphorical structure of the human conceptual system. *Cognitive Science, 4*, 195–208.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to Western thought*. Univeristy Chicago Press.

Lang, P. J. (1995). The Emotion Probe. *American Psychologist Association*, *50*(5), 372–385. https://doi.org/10.1037/0003-066X.50.5.372

Lenth, R. V. (2018). *emmeans: Estimated Marginal Means, aka Least-Squares Means.* (R package version 1.3.0).

Lester, A. W., Moffat, S. D., Wiener, J. M., Barnes, C. A., & Wolbers, T. (2017). The Aging Navigational System. *Neuron*, *95*(5), 1019–1035. https://doi.org/10.1016/j.neuron.2017.06.037

Lewis, M. B., & Bowler, P. J. (2008). Botulinum toxin cosmetic therapy correlates with a more positive mood. *Journal of Cosmetic Dermatology, 8*(1), 24–26. https://doi.org/https://doi.org/10.1111/j.1473-2165.2009.00419.x

Liepelt, R., Dolk, T., & Prinz, W. (2012). Bidirectional semantic interference between action and speech. *Psychological Research*, *76*, 446–455.

Liuzzi, A. G., Bruffaerts, R., & Vandenberghe, R. (2019). The medial temporal written word processing system. *Cortex*, *119*, 287–300. https://doi.org/10.1016/j.cortex.2019.05.002

Luce, P. A., & Pisoni, D. B. (1998). Recognizing Spoken Words: The Neighborhood Activation Model. *Ear & Hearing*, *19*, 1–36.

Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavioural Research, 49*, 1494–1502. https://doi.org/10.3758/s13428-016-0809-y

Ly, A., Verhagen, J., & Wagenmakers, E.-J. (2016). Harold Jeffreys's default Bayes factor hypothesis tests: Explanation, extension, and application in psychology. *Journal of Mathematical Psychology*, *72*, 19–32.

Majid, A., Bowerman, M., Kita, S., Haun, D. B. M., & Levinson, S. C. (2004). Can language restructure cognition? The case for space. *Trends in Cognitive Sciences*, *8*(3), 108–114. https://doi.org/10.1016/j.tics.2004.01.003

Marino, B. F. M., Gallese, V., Buccino, G., & Riggio, L. (2012). Language sensorimotor specificity modulates the motor system. *Cortex, 48*(7), 849–856. https://doi.org/10.1016/j.cortex.2010.12.003

Marino, B. F. M., Gough, P. M., Gallese, V., Riggio, L., & Buccino, G. (2013). How the motor system handles nouns: A behavioral study. *Psychological Research*, *77*(1). https://doi.org/10.1007/s00426-011-0371-2

Masson, M. E. J., Bub, D. N., & Warren, C. M. (2008). Kicking calculators: Contribution of embodied representations to sentence comprehension. *Journal of Memory and Language, 59*(3), 256–265. https://doi.org/10.1016/j.jml.2008.05.003

Meade, G. (2020). The role of phonology during visual word learning in adults: An integrative review. *Psychonomic Bulletin and Review, 27*(1), 15–23. https://doi.org/10.3758/s13423-019-01647-0

Migo, E. M., Mayes, A. R., & Montaldi, D. (2012). Measuring recollection and familiarity: Improving the remember/know procedure. *Consciousness and Cognition*, *21*(3), 1435–1455. https://doi.org/10.1016/j.concog.2012.04.014

Mirabella, G., Iaconelli, S., Spadacenta, S., Federico, P., & Gallese, V. (2012). Processing of hand-related verbs specifically affects the planning and execution of arm reaching movements. *PLoS ONE*, *7*(e35403).

Monsalves, M. J., Bangdiwala, A. S., Thabane, A., & Bangdiwala, S. I. (2020). LEVEL (Logical Explanations & Visualizations of Estimates in Linear mixed models): Recommendations for reporting multilevel data and analyses. *BMC Medical Research Methodology*, *20*(1), 1–9. https://doi.org/10.1186/s12874-019-0876-8

Montefinese, M., Ambrosini, E., Fairfield, B., & Mammarella, N. (2014). The adaptation of the affective norms for English words (ANEW) for Italian. *Behavior Research Methods*, *46*, 887–903.

Morey, R. D., Kaschak, M. P., Díez-Álamo, A. M., Glenberg, A. M., Zwaan, R. A., Lakens, D., Ibáñez, A., García, A., Gianelli, C., Jones, J. L., Madden, J., Alifano, F., Bergen, B., Bloxsom, N. G., Bub, D. N., Cai, Z. G., Chartier, C. R., Chatterjee, A., Conwell, E., … Noam Ziv-Crispel. (2022). A pre-registered, multi-lab non-replication of the Action-sentence Compatibility Effect (ACE). *Psychonomic Bulletin & Review*, *29*, 613–626.

Morganti, F. (2018). Enacting space in virtual reality: A comparison between Money's Road Map test and its virtual version. *Frontiers in Psychology*, *9*(DEC), 1–9. https://doi.org/10.3389/fpsyg.2018.02410

Nathan, G. S. (2017). Phonology. In *The Cambridge Handbook of Cognitive Linguistics* (pp. 214–228). Cambridge University Press.

Nazir, T. A., Jeannerod, M., & Hauk, O. (2008). The role of sensory-motor systems for language understanding. *Journal of Physiology*, *102*, 1–3.

Neuhoff, J. G. (2001). An Adaptive Bias in the Perception of Looming Auditory Motion.

Ecological Psychology, 13(2), 87–110. https://doi.org/10.1207/S15326969ECO1302_2

Neuhoff, J. G. (2018). Adaptive Biases in Visual and Auditory Looming Perception. In *Spatial Biases in Perception and Cognition* (pp. 180–190). Cambridge University Press. https://doi.org/10.1017/9781316651247.013

O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map . Preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, *34*, 171–175.

O'Keefe, J., & Nadel, L. (1978). *The Hippocampus as a Cognitive Map*. Clarendon Press.

Papeo, L., Corradi-Dell'Acqua, C., & Rumiati, R. I. (2011). "She" is not like "I": the tie between language and action is in our imagination. *Journal of Cognitive Neuroscience*, *23*, 3939–3948. https://doi.org/10.1162/jocn_a_00075

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–e141. https://doi.org/10.1037/xge0000125

Pecher, D., van Dantzig, S., Zwaan, R. A., & Zeelenberg, R. (2009). Language comprehenders retain implied shape and orientation of objects. *Journal of Experimental Psychology (Hove)*, *62*, 1108–1114.

Piaget, J. (1953). *The origins of intelligence in the child*. Routledge.

Piccardi, L., Guariglia, P., Nori, R., & Palmiero, M. (2020). The role of emotional landmarks in embodied and not-embodied tasks. *Brain Sciences*, *10*(2). https://doi.org/10.3390/brainsci10020058

Plancher, G., Tirard, A., Gyselinck, V., Nicolas, S., & Piolino, P. (2012). Using virtual reality to characterize episodic memory profiles in amnestic mild cognitive impairment and Alzheimer's disease: Influence of active and passive encoding. *Neuropsychologia*, *50*(5), 592–602. https://doi.org/10.1016/j.neuropsychologia.2011.12.013

R Core Team (2014). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. http://www.r-project.org/.%0A

Repetto, C., Colombo, B., Cipresso, P., & Riva, G. (2013). The effects of rTMS over the primary motor cortex: the link between action and language. *Neuropsychologia*, *51*(1), 8–13. https://doi.org/10.1016/j.neuropsychologia.2012.11.001S0028-3932(12)00468-X [pii]

Repetto, C., Mathias, B., Weichselbaum, O., & Macedonia, M. (2021). Visual recognition of words learned with gestures induces motor resonance in the forearm muscles. *Scientific Reports*, *11*(1), 1–16. https://doi.org/10.1038/s41598-021-96792-9

Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, *6*(2), 135–147. https://doi.org/10.1016/j.edurev.2010.12.001

Riva, G., Castelnuovo, G., & Mantovani, F. (2006). Transformation of flow in rehabilitation: the role of advanced communication technologies. *Behavior Research Methods*, *38*(2), 237–244. https://doi.org/10.3758/BF03192775

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review OfNeuroscience*, *27*, 169–192.

Ruotolo, F., Claessen, M. H. G., & van der Ham, I. J. M. (2018). Putting emotions in routes: the influence of emotionally laden landmarks on spatial memory. *Psychological Research*, *83*(5), 1083–1095. https://doi.org/10.1007/s00426-018-1015-6

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*, 83–90.

Schneider, I. K., & Mattes, A. (2021). The effect of spatial distance between objects on categorization level. *Psychonomic Bulletin and Review*, 230–242. https://doi.org/10.3758/s13423-021-01949-2

Schneider, I. K., Rutjens, B. T., Jostmann, N. B., & Lakens, D. (2011). Weighty Matters: Importance Literally Feels Heavy. *Social Psychological and Personality Science*, *2*(5), 474–478. https://doi.org/10.1177/1948550610397895

Schütz-Bosbach, S., Mancini, B., Aglioti, S. M., & Haggard, P. (2006). Self and Other in the

Human Motor System. *Current Biology*, *16*(18), 1830–1834. https://doi.org/10.1016/j.cub.2006.07.048

Serino, S., Cipresso, P., Morganti, F., & Riva, G. (2014). The role of egocentric and allocentric abilities in Alzheimer's disease: A systematic review. *Ageing Research Reviews*, *16*, 32–44. https://doi.org/10.1016/j.arr.2014.04.004

Serratrice, L., Sorace, A., Filiaci, F., & Baldo, M. (2011). Pronominal objects in English-Italian and Spanish-Italian bilingual children. *Applied Psycholinguistics*, *27*, 1–27. https://doi.org/10.1017/S0142716411000543

Shunji, A. (2011). The action-sentence compatibility effect in Japanese sentences. *Perceptual and Motor Skills*, *113*(2), 597–604. https://doi.org/10.2466/28.PMS.113.5.597-604

Sinha, C., & Kristine, J. D. L. (2001). Language, Culture and the Embodiment of Spatial Cognition. *Cognitive Linguistics*, *11*(1–2), 17–41. https://doi.org/10.1515/cogl.2001.008

Solomon, E. A., Lega, B. C., Sperling, M. R., & Kahana, M. J. (2019). Hippocampal theta codes for distances in semantic and temporal spaces. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, *116*(48), 24343–24352. https://doi.org/10.1073/pnas.1906729116

Spreng, R. N., Mar, R. A., & Kim, A. S. N. (2008). The Common Neural Basis of Autobiographical Memory, Prospection, Navigation, Theory of Mind, and the Default Mode: A Quantitative Meta-analysis. *Journal of Cognitive Neuroscience*, *21*(3), 489–510. https://doi.org/10.1162/jocn.2008.21029

Squire, L. R., Wixted, J., & R, C. (2007). Recognition memory and the medial temporal lobe: a new perspective. *Nature Reviews. Neuroscience*, *8*(11), 872–883. https://doi.org/10.1038/nrn2154

Stamer, M. K., & Vitevitch, M. S. (2012). Phonological similarity influences word learning in adults learning Spanish as a foreign language. *Bilingualism*, *15*(3), 490–502. https://doi.org/10.1017/S1366728911000216

Steel, A., Robertson, C. E., & Taube, J. S. (2021). Current Promises and Limitations of Combined

Virtual Reality and Functional Magnetic Resonance Imaging Research in Humans: A Commentary on Huffman and Ekstrom (2019). *Journal of Cognitive Neuroscience*, *33*(2), 159–166. https://doi.org/10.1162/jocn_a_01635

Sui, J., & Humphreys, G. W. (2015). The Integrative Self: How Self-Reference Integrates Perception and Memory. *Trends in Cognitive Sciences*, *19*(12), 719–728. https://doi.org/10.1016/j.tics.2015.08.015

Sullivan, G. M., & Feinn, R. (2012). Using Effect Size-or Why the P Value Is Not Enough. *Journal of Graduate Medical Education*, *4*(3), 279–282.

Talmy, L. (1983). How language structures space. In H. Pick & L. Acredelo (Eds.), *Spatial Orientation: Theory, Research, and Application* (pp. 225–282). Plenum.

Tardif, T., & Shatz, M. (1997). Caregiver speech and children’s use of nouns versus verbs : A comparison of English, Italian, and Mandarin*. *Journal of Child Language*, *24*, 535–565.

Taube, J. S., Valerio, S., & Yoder, R. M. (2013). Is Navigation in Virtual Reality with fMRI Really Navigation? *Journal of Cognitive Neuroscience*, *25*(7), 1008–1019. https://doi.org/10.1162/jocn

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. https://doi.org/10.1162/0898929053124965

Tomasino, B., Werner, C. J., Weiss, P. H., & Fink, G. R. (2007). Stimulus properties matter more than perspective: An fMRI study of mental imagery and silent reading of action phrases. *NeuroImage*, *36*(SUPPL. 2), 128–141. https://doi.org/10.1016/j.neuroimage.2007.03.035

Tuena, C., Mancuso, V., Stramba-Badiale, C., Pedroli, E., Stramba-Badiale, M., Riva, G., & Repetto, C. (2021). Egocentric and Allocentric Spatial Memory in Mild Cognitive Impairment with Real-World and Virtual Navigation Tasks: A Systematic Review. *Journal of Alzheimer’s Disease*, *79*(1), 95–116. https://doi.org/10.3233/JAD-201017

Tuena, C., Serino, S., Dutriaux, L., Riva, G., & Piolino, P. (2019). Virtual Enactment Effect on Memory in Young and Aged Populations : a Systematic Review. *Journal of Clinical Medicine*, *8*(5), 620. https://doi.org/10.3390/jcm8050620

Tuena, C., Serino, S., Gaston-Bellegarde, A., Orriols, E., Makowski, D., Riva, G., & Piolino, P. (2017). How virtual embodiment affects episodic memory functioning : a proof-of-concept study. *Annual Review of Cybertherapy and Telemedicine*, 98–103.

Tuena, C., Serino, S., Pedroli, E., Stramba-Badiale, M., Riva, G., & Repetto, C. (2021). Building Embodied Spaces for Spatial Memory Neurorehabilitation with Virtual Reality in Normal and Pathological Aging. *Brain Sciences*, *11*(8), 1067.

Tulving, E. (2001). Episodic memory and common sense: how far apart? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *356*(1413), 1505–1515. https://doi.org/10.1098/rstb.2001.0937

Tulving, E., & Markowitsch, H. J. (1998). Episodic and declarative memory: Role of the hippocampus. *Hippocampus*, *8*(3), 198–204. https://doi.org/10.1002/(SICI)1098-1063(1998)8:3<198::AID-HIPO2>3.0.CO;2-G

Tversky, B., & Hard, B. M. (2009). Embodied and disembodied cognition: Spatial perspective-taking. *Cognition*, *110*(1), 124–129. https://doi.org/10.1016/j.cognition.2008.10.008

Vitevitch, M. S., & Rodríguez, E. (2005). Neighborhood density effects in spoken word recognition in Spanish. *Journal of Multilingual Communication Disorders*, *3*(1), 64–73. https://doi.org/10.1080/14769670400027332

Vogeley, K., & Fink, G. R. (2003). Neural correlates of the first-person-perspective. *Trends in Cognitive Sciences*, *7*(1), 38–42. https://doi.org/10.1016/S1364-6613(02)00003-7

Vogeley, K., May, M., Ritzl, a, Falkai, P., Zilles, K., & Fink, G. R. (2004). Neural correlates of first-person perspective as one constituent of human self-consciousness. *Journal of Cognitive Neuroscience*, *16*(5), 817–827. https://doi.org/10.1162/089892904970799

Vukovic, N., & Shtyrov, Y. (2017). Cortical networks for reference-frame processing are shared by language and spatial navigation systems. *NeuroImage, 161*(January), 120–133. https://doi.org/10.1016/j.neuroimage.2017.08.041

Vukovic, N., & Williams, J. N. (2015). Individual differences in spatial cognition influence mental simulation of language. *Cognition*, *142*, 110–122. https://doi.org/10.1016/j.cognition.2015.05.017

Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General*, *143*(5), 2020–2045.

Willems, R. M., Hagoort, P., & Casasanto, D. (2010). Body-specific representations of action verbs: Neural evidence from right- and left-handers. *Psychological Science*, *21*(1), 67–74. https://doi.org/10.1177/0956797609354072

Wilson, A. D., & Golonka, S. (2013). Embodied Cognition is Not What you Think it is. *Frontiers in Psychology*, *4*(58). https://doi.org/10.3389/fpsyg.2013.00058

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*(4), 625–636. https://doi.org/10.3758/BF03196322

Wixted, J. T., & Stretch, V. (2004). In defense of the signal detection interpretation of remember/know judgments. *Psychonomic Bulletin and Review*, *11*(4), 616–641. https://doi.org/10.3758/BF03196616

Wolbers, T., Klatzky, R. L., Loomis, J. M., Wutte, M. G., & Giudice, N. A. (2011). Modality-independent coding of spatial layout in the human brain. *Current Biology, 21*(11), 984–989. https://doi.org/10.1016/j.cub.2011.04.038

Wu, Z., & Palmer, M. (1994). Verbs semantics and lexical selection. *Proceedings of the 32nd Annual Meeting on Association for Computational Linguistics*, 133–138.

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science*, *6*(5), 292–297.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation Models in Language Comprehension and

Memory. *Psychological Bulletin*, *123*(2), 162–185.

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, *13*(2), 168–171. https://doi.org/10.1111/1467-9280.00430

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

Zwaan, R. A., van der Stoep, N., Guadalupe, T., & Bouwmeester, S. (2012). Language Comprehension in the Balance: The Robustness of the Action-Compatibility Effect (ACE). *PLoS One*, *7*(2), e31204.

Appendix

A

Sentence	**Direction**	**Type**
You told Liz the story.	Away	abstract
You taught him a lesson.	Away	abstract
You sent your family regards.	Away	abstract
You pitched Andy the idea.	Away	abstract
You blew your sister a kiss.	Away	abstract
You offered Chris some writing tips.	Away	abstract
You gave Jesse another chance.	Away	abstract
You paid Amanda tribute.	Away	abstract
You sang Jenni a song.	Away	abstract
You lavished Steve with praise.	Away	abstract
You devoted your time to Tiana.	Away	abstract
You received the complaint from Ian.	Away	abstract
You transferred responsibility to Anna.	Away	abstract
You confessed your secret to Dan.	Away	abstract
You dedicated the song to John.	Away	abstract
Dan confessed his secret to you.	away	abstract
You bestowed the honor upon Art.	Away	abstract
Mike sold the land to you.	Away	abstract
You conveyed the message to Adam.	Away	abstract
You radioed the message to the policeman.	Away	abstract
You transmitted the orders to Sara.	Away	abstract
Your family sent you regards.	Toward	abstract
He taught you a lesson.	Toward	abstract
Liz told you the story.	Toward	abstract
Your sister blew you a kiss.	Toward	abstract
Andy pitched you the idea.	Toward	abstract

Jenni sang you a song.	Toward	abstract
Steve lavished you with praise.	Toward	abstract
Jesse gave you another chance.	Toward	abstract
Amanda paid you tribute.	Toward	abstract
Chris offered you some writing tips.	Toward	abstract
Anna transferred responsibility to you.	Toward	abstract
Ian received the complaint from you.	Toward	abstract
Tiana devoted her time to you.	Toward	abstract
John dedicated the song to you.	Toward	abstract
The policeman radioed the message to you.	Toward	abstract
Art bestowed the honor upon you.	Toward	abstract
Adam conveyed the message to you.	Toward	abstract
You sold the land to Mike.	Toward	abstract
Sara transmitted the orders to you.	Toward	abstract
You poured your dad some water.	Away	concrete
You hit Paul the baseball.	Away	concrete
You shot Shawn the rubber band.	Away	concrete
You slipped Heather a note.	Away	concrete
You rolled Mike the marble.	Away	concrete
You kicked Joe the soccer ball.	Away	concrete
You threw Diane the pen.	Away	concrete
You bought Christine ice cream.	Away	concrete
You slid Sally the cafeteria tray.	Away	concrete
You handed Courtney the notebook.	Away	concrete
You handed the puppy to Katie.	Away	concrete
You dealt the cards to Mark.	Away	concrete
You entrusted the key to Jeff.	Away	concrete
You drove the car to Amber.	Away	concrete
You dispensed the rations to Kelly.	Away	concrete
You donated money to Vincent.	Away	concrete

You delivered the pizza to Andy.	Away	concrete
You forked over the cash to Alex.	Away	concrete
You kicked the football to Jack.	Away	concrete
You awarded a medal to Helen.	Away	concrete
Shawn shot you the rubber band.	Toward	concrete
Heather slipped you a note.	Toward	concrete
Your dad poured you some water.	Toward	concrete
Paul hit you the baseball.	Toward	concrete
Mike rolled you the marble.	Toward	concrete
Diane threw you the pen.	Toward	concrete
Sally slid you the cafeteria tray.	Toward	concrete
Christine bought you ice cream.	Toward	concrete
Courtney handed you the notebook.	Toward	concrete
Joe kicked you the soccer ball.	Toward	concrete
Mark dealt the cards to you.	Toward	concrete
Kelly dispensed the rations to you.	Toward	concrete
Katie handed the puppy to you.	Toward	concrete
Jeff entrusted the key to you.	Toward	concrete
Amber drove the car to you.	Toward	concrete
Alex forked over the cash to you.	Toward	concrete
Helen awarded a medal to you.	Toward	concrete
Andy delivered the pizza to you.	Toward	concrete
Vincent donated money to you.	Toward	concrete
Jack kicked the football to you.	Toward	concrete
Howard joked the ice cream to you.	-	nonsense
Frank laundered the bench to you.	-	nonsense
You cleaned the pizza to Jon.	-	nonsense
Ashley locked the lint roller to you.	-	nonsense
You floundered the train to Daniel.	-	nonsense
You gargled integrity to Nicole.	-	nonsense

Charles bruised you the hamburger.	-	nonsense
You carpeted directions to Emily.	-	nonsense
You forked the bottle to Lori.	-	nonsense
Donald drank the shovel with you.	-	nonsense
Edward flew the house with you.	-	nonsense
You mowed Nick your opinion.	-	nonsense
You snored justice for Scott.	-	nonsense
You sang the pizza to Jack.	-	nonsense
Julie broke the writing tips on you.	-	nonsense
Anne mingled the complaint to you.	-	nonsense
You licked the jacket to Paul.	-	nonsense
You sunk Lillian your viewpoint.	-	nonsense
You ingested the car with Ben.	-	nonsense
Craig ingested you instructions.	-	nonsense
You cleaned the honor for Debra.	-	nonsense
Debbie posted you the flowers.	-	nonsense
You sang the marble with Katherine.	-	nonsense
Gerald shambled you loyalty.	-	nonsense
You cleaned responsibility with Jane.	-	nonsense
Rose parted you the trailer.	-	nonsense
You flexed Joyce a moment.	-	nonsense
You smelled the song with Irene.	-	nonsense
You laughed the pen to Ken.	-	nonsense
Susan rehearsed the medal to you.	-	nonsense
Stephanie frowned the door for you.	-	nonsense
Michael bit the message to you.	-	nonsense
Patrick blanketed you the chance.	-	nonsense
You parked the memo to Sandra.	-	nonsense
Justin tossed you with the paintball.	-	nonsense
Martha parked the string to you.	-	nonsense

You sneezed Thomas secrets.	-	nonsense
You hanged honesty with Laura.	-	nonsense
Roy tasted the soccer ball to you.	-	nonsense
You danced the land to Chris.	-	nonsense
Bonnie smiled the key to you.	-	nonsense
Rachel festered relief to you.	-	nonsense
You heated Annie the blame.	-	nonsense
You drank your idea to Carol.	-	nonsense
You drank the baseball with Eugene.	-	nonsense
You bled the rations to Joe.	-	nonsense
You sunk Norma the monitor.	-	nonsense
Jeffery washed you the thought.	-	nonsense
You glued the story with Donna.	-	nonsense
Ronald scratched the hat to you.	-	nonsense
Terry medicated commands to you.	-	nonsense
You radioed the floor with Louis.	-	nonsense
You blew a car to Frances.	-	nonsense
You cleaned the puppy to Clarence.	-	nonsense
Marie ate the regards to you.	-	nonsense
You cooked Joan duties.	-	nonsense
You tasted the papers to Ronald.	-	nonsense
You swam Douglas the truth.	-	nonsense
Lisa trudged you the concept.	-	nonsense
You tickled the orders with Raymond.	-	nonsense
You barked the football to Martin.	-	nonsense
You held the chance to Tim.	-	nonsense
Joshua snored the frame with you.	-	nonsense
Beverly hanged you more time.	-	nonsense
Melissa flushed you the appeal.	-	nonsense
George rolled you adoration.	-	nonsense

Cheryl jousted you the marker.	-	nonsense
You harpooned Betty the sheet.	-	nonsense
You perfumed Steven accolades.	-	nonsense
Andrew paddled fairness to you.	-	nonsense
You kissed the time to Paula.	-	nonsense
Jesse pickled praise on you.	-	nonsense
Phyllis cleaned honor upon you.	-	nonsense
You loafed the coffee cup to Peter.	-	nonsense
Brandon choked the lesson with you.	-	nonsense
You retaliated Arthur the opportunity.	-	nonsense
You flew on the note to Judith.	-	nonsense
You bordered Brian the chain.	-	nonsense
Henry wedged an homage to you.	-	nonsense
You fell the message to Judy.	-	nonsense

B

Sentence	**Direction**	**Type**
Luca mi sta tirando un elastico	Toward	concrete
Sto tirando l'elastico a Sofia	Away	concrete
Sara mi sta versando dell'acqua	Toward	concrete
Sto versando dell'acqua a mio padre	Away	concrete
Marco mi sta colpendo con la palla	Toward	concrete
Sto colpendo Mario con la palla	Away	concrete
Mario mi sta passando un bigliettino	Toward	concrete
Sto passando un bigliettino a Mario	Away	concrete
Carlo sta facendo rotolare la biglia verso di me	Toward	concrete
Sto facendo rotolare la biglia verso Sandro	Away	concrete
Emma sta scagliando la palla di neve contro di me	Toward	concrete
Sto scagliando la palla di neve contro Nadia	Away	concrete
Giulia sta gettando la penna verso di me	Toward	concrete
Sto gettando la penna verso Laura	Away	concrete
Paolo sta facendo scivolare il vassoio verso di me	Toward	concrete
Sto facendo scivolare il vassoio verso Marco	Away	concrete
Chiara mi sta dando un gelato	Toward	concrete
Sto dando un gelato a Claudia	Away	concrete
Claudio mi sta consegnando il quaderno	Toward	concrete
Sto consegnando il quaderno a Sara	Away	concrete
Livio mi sta calciando la palla	Toward	concrete
Sto calciando la palla a Luca	Away	concrete
Claudio mi sta passando il cucciolo	Toward	concrete
Sto passando il cucciolo ad Anna	Away	concrete
Bruno mi sta dando le carte	Toward	concrete
Sto dando le carte a Laura	Away	concrete
Maria mi sta affidando le chiavi	Toward	concrete

Sto affidando le chiavi a Sandra	Away	concrete
Laura mi sta distribuendo le razioni	Toward	concrete
Sto distribuendo le razioni a Dario	Away	concrete
Paola mi sta portando l'automobile	Toward	concrete
Sto portando l'automobile ad Alessia	Away	concrete
Fulvio mi sta sganciando i soldi	Toward	concrete
Sto sganciando i soldi a Marco	Away	concrete
Anna mi sta donando del denaro	Toward	concrete
Sto donando del denaro a Laura	Away	concrete
Alessia mi sta premiando con una medaglia	Toward	concrete
Sto premiando Sonia con una medaglia	Away	concrete
Bruno mi sta consegnando la pizza	Toward	concrete
Sto consegnando la pizza a Dario	Away	concrete
Marco mi sta scherzando il gelato	-	nonsense
Franco sta cuocendo la panca per me	-	nonsense
Sto pulendo la pizza per nome	-	nonsense
Lisa mi sta chiudendo il vaso	-	nonsense
Sto impastando il treno a Lucio	-	nonsense
Sto nuotando la morale a Nicola	-	nonsense
Sandra mi sta sbucciando l'hamburger	-	nonsense
Sto rovesciando la direzione a Luca	-	nonsense
Sto sparando la bottiglia per Sara	-	nonsense
Giulia sta bevendo la pala con me	-	nonsense
Laura sta volando la casa su di me	-	nonsense
Sto colorando ad Alessia la mia opinione	-	nonsense
Sto rompendo la giustizia a Luca	-	nonsense
Sto cantando la pizza a Marco	-	nonsense
Marco mi sta rompendo i pareri	-	nonsense
Laura mi sta mischiando la denuncia	-	nonsense
Sto leccando verità a Davide	-	nonsense

Sto distruggendo a Livia la mia visuale	-	nonsense
Sto mangiando l'auto con Bruno	-	nonsense
Mario sta ingerendo le colpe	-	nonsense
Sto pulendo l'onore per Livio	-	nonsense
Sara mi sta imbucando dei fiori	-	nonsense
Sto cantando il marmo con Paola	-	nonsense
Lidia mi sta trascinando fedeltà	-	nonsense
Sto sorvolando la responsabilità con Nadia	-	nonsense
Carla mi sta dividendo il telefono	-	nonsense
Sto piegando a Mario un attimo	-	nonsense
Sto odorando la canzone con Irene	-	nonsense
Sto ridendo la penna a Mauro	-	nonsense
Marco mi sta provando la medaglia	-	nonsense
Nadia sta assaggiando la porta per me	-	nonsense
Carlo mi sta mordendo il messaggio	-	nonsense
Sandro mi sta ricoprendo la possibilità	-	nonsense
Sto parcheggiando il promemoria per Lidia	-	nonsense
Sara mi sta gettando con il fango	-	nonsense
Luca mi sta parcheggiando la corda	-	nonsense
Sto starnutendo i segreti di Maria	-	nonsense
Sto appendendo l'onestà con Laura	-	nonsense
Silvia mi sta intervallando la sfera	-	nonsense
Sto danzando la terra a Teresa	-	nonsense

C

List Sentence	Perspective
Luca mi sta tirando un elastico	Third-person
Sto tirando l'elastico a Sofia	First-person
Sara mi sta versando l'acqua	Third-person
Sto versando dell'acqua a mio padre	First-person
Marco mi sta porgendo la palla da calcio	Third-person
Sto porgendo a Mario la palla da calcio	First-person
Mario mi sta passando un bigliettino	Third-person
Sto passando un bigliettino a Mario	First-person
Carlo mi sta restituendo la biglia	Third-person
Sto restituendo la biglia a Sandro	First-person
Emma mi sta regalando la palla di carta	Third-person
Sto regalando la palla di carta a Nadia	First-person
Giulia mi sta prestando la penna	Third-person
Sto prestando la penna a Laura	First-person
Paolo mi sta porgendo il vassoio	Third-person
Sto porgendo il vassoio a Marco	First-person
Chiara mi sta dando un gelato	Third-person
Sto dando un gelato a Claudia	First-person
Claudio mi sta consegnando il quaderno	Third-person
Sto consegnando il quaderno a Sara	First-person
Livio mi sta passando la palla	Third-person
Sto passando la palla a Luca	First-person
Claudio mi sta passando il pupazzo	Third-person
Sto passando il pupazzo ad Anna	First-person
Bruno mi sta dando le carte	Third-person
Sto dando le carte a Laura	First-person
Maria mi sta affidando le chiavi	Third-person
Sto affidando le chiavi a Sandra	First-person
Luca mi sta distribuendo le razioni	Third-person
Sto distribuendo le razioni a Dario	First-person
Paola mi sta rendendo la macchinina	Third-person
Sto rendendo la macchinina ad Alessia	First-person
Fulvio mi sta sganciando i soldi	Third-person
Sto sganciando i soldi a Marco	First-person
Anna mi sta donando del denaro	Third-person
Sto donando del denaro a Laura	First-person

Alessia mi sta premiando con una medaglia	Third-person
Sto premiando Sonia con una medaglia	First-person
Bruno mi sta consegnando la pizza	Third-person
Sto consegnando la pizza a Dario	First-person

D

List Sentence	Perspective
Alex is shooting the rubber band at me	Third-person
I am shooting the rubber band at Luke	First-person
Sarah is pouring me a cup of water	Third-person
I am pouring a cup of water for my father	First-person
Marc is passing me the soccer ball	Third-person
I am passing the soccer ball to Marc	First-person
Hugo is passing me a note	Third-person
I am passing a note to Daniel	First-person
Charles is giving me back the marble	Third-person
I am giving the marble back to Alex	First-person
Emma is giving me a paper ball	Third-person
I am giving a paper ball to Ella	First-person
Mary is lending me a pen	Third-person
I am lending a pen to Jasmine	First-person
Paul is passing me the tray	Third-person
I am passing the tray to Marc	First-person
Claire is giving me ice-cream	Third-person
I am giving ice-cream to Celine	First-person
Claire is handing me the notebook	Third-person
I am handing the notebook to Sarah	First-person
Matthew is passing me the ball	Third-person
I am passing the ball to Luke	First-person
Charles is passing me the puppet	Third-person
I am passing the puppet to Hanna	First-person
Jack is dealing me the cards	Third-person
I am dealing the cards to Peter	First-person
Mary is handing me the keys	Third-person
I am handing the keys to Sandra	First-person
Laura is breaking off a piece of chocolate for me	Third-person
I am breaking off a piece of chocolate for Bill	First-person
Paula is handing me the little car	Third-person
I am handing the little car to Alexia	First-person
Bill is giving me the money	Third-person
I am giving the money to Marc	First-person
Hanna is donating money to me	Third-person
I am donating money to Laura	First-person
Alexia is awarding me a medal	Third-person

I am awarding Bobbie a medal	First-person
Irwin is delivering a pizza to me	Third-person
I am delivering a pizza to Bill	First-person

E

frase	perspective	type	target
ENCODING BLOCK 1			
Gianni ti scaglia l'elastico	3	action	OLD
Tu lanci il libro a Gianni	1	action	OLD
Gianni ti versa il vino	3	action	OLD
Tu mesci la birra a Gianni	1	action	OLD
Gianni ti tira la biro	3	action	OLD
Tu calci la palla a Gianni	1	action	OLD
Gianni ti porge un foglio	3	action	OLD
Tu presenti un biglietto a Gianni	1	action	OLD
Gianni ti restituisce la biglia	3	action	OLD
Tu doni la bici a Gianni	1	action	OLD
Gianni ti regala il quadro	3	action	OLD
Tu allunghi il bicchiere a Gianni	1	action	OLD
Gianni ti presta la penna	3	action	OLD
Tu cedi la matita a Gianni	1	action	OLD
Gianni ti butta una moneta	3	action	OLD
Tu offri il vassoio a Gianni	1	action	OLD
Gianni ti invia il pacco	3	action	OLD
Tu favorisci i documenti a Gianni	1	action	OLD
Gianni ti recapita la lettera	3	action	OLD
Tu ridai gli occhiali a Gianni	1	action	OLD
Gianni ti spiega la lezione di storia	3	abstract	OLD
Tu discuti delle responsabilità con Gianni	1	abstract	OLD
Gianni ti chiede un consiglio	3	abstract	OLD
Tu escogiti una soluzione a Gianni	1	abstract	OLD
Gianni ti omette delle informazioni	3	abstract	OLD
Tu impartisci delle regole a Gianni	1	abstract	OLD
Gianni ti dimostra fiducia	3	abstract	OLD
Tu infondi serenità a Gianni	1	abstract	OLD
Gianni ti gestisce le finanze	3	abstract	OLD
Tu soddisfi una richiesta a Gianni	1	abstract	OLD
Gianni ti propone un acquisto	3	abstract	OLD
Tu domandi un parere a Gianni	1	abstract	OLD
Gianni ti concede del tempo	3	abstract	OLD
Tu introduci un argomento a Gianni	1	abstract	OLD
Gianni ti palesa affetto	3	abstract	OLD
Tu aumenti i vantaggi a Gianni	1	abstract	OLD
Gianni ti porta rispetto	3	abstract	OLD
Tu consenti una pausa a Gianni	1	abstract	OLD
Gianni ti influenza l'opinione	3	abstract	OLD

Tu risolvi problemi a Gianni	1	abstract	OLD
RECOGNITION BLOCK 1			
Gianni ti scaglia l'elastico	3	action	OLD
Tu lanci il libro a Gianni	1	action	OLD
Gianni ti versa il vino	3	action	OLD
Tu mesci la birra a Gianni	1	action	OLD
Gianni ti tira la biro	3	action	OLD
Tu calci la palla a Gianni	1	action	OLD
Gianni ti porge un foglio	3	action	OLD
Tu presenti un biglietto a Gianni	1	action	OLD
Gianni ti restituisce la biglia	3	action	OLD
Tu doni la bici a Gianni	1	action	OLD
Gianni ti regala il quadro	3	action	OLD
Tu allunghi il bicchiere a Gianni	1	action	OLD
Gianni ti presta la penna	3	action	OLD
Tu cedi la matita a Gianni	1	action	OLD
Gianni ti butta una moneta	3	action	OLD
Tu offri il vassoio a Gianni	1	action	OLD
Gianni ti invia il pacco	3	action	OLD
Tu favorisci i documenti a Gianni	1	action	OLD
Gianni ti recapita la lettera	3	action	OLD
Tu ridai gli occhiali a Gianni	1	action	OLD
Gianni ti spiega la lezione di storia	3	abstract	OLD
Tu discuti delle responsabilità con Gianni	1	abstract	OLD
Gianni ti chiede un consiglio	3	abstract	OLD
Tu escogiti una soluzione a Gianni	1	abstract	OLD
Gianni ti omette delle informazioni	3	abstract	OLD
Tu impartisci delle regole a Gianni	1	abstract	OLD
Gianni ti dimostra fiducia	3	abstract	OLD
Tu infondi serenità a Gianni	1	abstract	OLD
Gianni ti gestisce le finanze	3	abstract	OLD
Tu soddisfi una richiesta a Gianni	1	abstract	OLD
Gianni ti propone un acquisto	3	abstract	OLD
Tu domandi un parere a Gianni	1	abstract	OLD
Gianni ti concede del tempo	3	abstract	OLD
Tu introduci un argomento a Gianni	1	abstract	OLD
Gianni ti palesa affetto	3	abstract	OLD
Tu aumenti i vantaggi a Gianni	1	abstract	OLD
Gianni ti porta rispetto	3	abstract	OLD
Tu consenti una pausa a Gianni	1	abstract	OLD
Gianni ti influenza l'opinione	3	abstract	OLD
Tu risolvi problemi a Gianni	1	abstract	OLD
Tu scagli l'elastico a Gianni	1	action	NEW

Gianni ti lancia il libro	3	action	NEW
Tu versi il vino a Gianni	1	action	NEW
Gianni ti mesce la birra	3	action	NEW
Tu tiri la biro a Gianni	1	action	NEW
Gianni ti calcia la palla	3	action	NEW
Tu porgi un foglio a Gianni	1	action	NEW
Gianni ti presenta un biglietto	3	action	NEW
Tu restituisci la biglia a Gianni	1	action	NEW
Gianni ti dona la bici	3	action	NEW
Tu regali il quadro a Gianni	1	action	NEW
Gianni ti allunga il bicchiere	3	action	NEW
Tu presti la penna a Gianni	1	action	NEW
Gianni ti cede la matita	3	action	NEW
Tu butti una moneta a Gianni	1	action	NEW
Gianni ti offre il vassoio	3	action	NEW
Tu invii il pacco a Gianni	1	action	NEW
Gianni ti favorisce i documenti	3	action	NEW
Tu recapiti la lettera a Gianni	1	action	NEW
Gianni ti ridà gli occhiali	3	action	NEW
Tu spieghi la lezione di storia a Gianni	1	abstract	NEW
Gianni discute con te delle responsabilità	3	abstract	NEW
Tu chiedi un consiglio a Gianni	1	abstract	NEW
Gianni ti escogita una soluzione	3	abstract	NEW
Tu ometti delle informazioni a Gianni	1	abstract	NEW
Gianni ti impartisce delle regole	3	abstract	NEW
Tu dimostri fiducia a Gianni	1	abstract	NEW
Gianni ti infonde serenità	3	abstract	NEW
Tu gestisci le finanze a Gianni	1	abstract	NEW
Gianni ti soddisfa una richiesta	3	abstract	NEW
Tu proponi un acquisto a Gianni	1	abstract	NEW
Gianni ti domanda un parere	3	abstract	NEW
Tu concedi del tempo a Gianni	1	abstract	NEW
Gianni ti introduce un argomento	3	abstract	NEW
Tu palesi affetto a Gianni	1	abstract	NEW
Gianni ti aumenta i vantaggi	3	abstract	NEW
Tu porti rispetto a Gianni	1	abstract	NEW
Gianni ti consente una pausa	3	abstract	NEW
Tu influenzi l'opinione a Gianni	1	abstract	NEW
Gianni ti risolve problemi	3	abstract	NEW
ENCODING BLOCK 2			
Maria ti scaraventa la boccia	3	action	OLD
Tu getti il tappo a Maria	1	action	OLD
Maria ti passa il gioco	3	action	OLD

Tu impresti il phon a Maria	1	action	OLD
Maria ti da le carte	3	action	OLD
Tu assegni le mappe a Maria	1	action	OLD
Maria ti affida le chiavi	3	action	OLD
Tu affibbi il cane a Maria	1	action	OLD
Maria ti distribuisce le razioni	3	action	OLD
Tu trasferisci le foto a Maria	1	action	OLD
Maria ti riconsegna l'auto	3	action	OLD
Tu rendi la moto a Maria	1	action	OLD
Maria ti sgancia lo stipendio	3	action	OLD
Tu sborsi i soldi a Maria	1	action	OLD
Maria ti elargisce denaro	3	action	OLD
Tu fornisci l'ombrello a Maria	1	action	OLD
Maria ti premia con una coppa	3	action	OLD
Tu omaggi con un regalo Maria	1	action	OLD
Maria ti consegna il farmaco	3	action	OLD
Tu somministri la pillola a Maria	1	action	OLD
Maria ti suggerisce un romanzo	3	abstract	OLD
Tu fai alcuni favori a Maria	1	abstract	OLD
Maria ti espone dei dubbi	3	abstract	OLD
Tu nascondi un segreto a Maria	1	abstract	OLD
Maria ti consiglia una vacanza	3	abstract	OLD
Tu proteggi gli averi a Maria	1	abstract	OLD
Maria ti descrive la situazione	3	abstract	OLD
Tu valuti un affare con Maria	1	abstract	OLD
Maria ti insegna l'inglese	3	abstract	OLD
Tu ribadisci un'idea a Maria	1	abstract	OLD
Maria ti autorizza il trasferimento	3	abstract	OLD
Tu raddoppi il lavoro a Maria	1	abstract	OLD
Maria ti avverte del pericolo	3	abstract	OLD
Tu ricordi un evento a Maria	1	abstract	OLD
Maria ti stravolge il pensiero	3	abstract	OLD
Tu provochi una reazione a Maria	1	abstract	OLD
Maria ti esorta allo studio	3	abstract	OLD
Tu riconduci un'azione a Maria	1	abstract	OLD
Maria ti impone una prassi	3	abstract	OLD
Tu progetti una gita con Maria	1	abstract	OLD
RECOGNITION BLOCK 2			
Maria ti scaraventa la boccia	3	action	OLD
Tu getti il tappo a Maria	1	action	OLD
Maria ti passa il gioco	3	action	OLD
Tu impresti il phon a Maria	1	action	OLD
Maria ti da le carte	3	action	OLD

Tu assegni le mappe a Maria	1	action	OLD
Maria ti affida le chiavi	3	action	OLD
Tu affibbi il cane a Maria	1	action	OLD
Maria ti distribuisce le razioni	3	action	OLD
Tu trasferisci le foto a Maria	1	action	OLD
Maria ti riconsegna l'auto	3	action	OLD
Tu rendi la moto a Maria	1	action	OLD
Maria ti sgancia lo stipendio	3	action	OLD
Tu sborsi i soldi a Maria	1	action	OLD
Maria ti elargisce denaro	3	action	OLD
Tu fornisci l'ombrello a Maria	1	action	OLD
Maria ti premia con una coppa	3	action	OLD
Tu omaggi con un regalo Maria	1	action	OLD
Maria ti consegna il farmaco	3	action	OLD
Tu somministri la pillola a Maria	1	action	OLD
Maria ti suggerisce un romanzo	3	abstract	OLD
Tu fai alcuni favori a Maria	1	abstract	OLD
Maria ti espone dei dubbi	3	abstract	OLD
Tu nascondi un segreto a Maria	1	abstract	OLD
Maria ti consiglia una vacanza	3	abstract	OLD
Tu proteggi gli averi a Maria	1	abstract	OLD
Maria ti descrive la situazione	3	abstract	OLD
Tu valuti un affare con Maria	1	abstract	OLD
Maria ti insegna l'inglese	3	abstract	OLD
Tu ribadisci un'idea a Maria	1	abstract	OLD
Maria ti autorizza il trasferimento	3	abstract	OLD
Tu raddoppi il lavoro a Maria	1	abstract	OLD
Maria ti avverte del pericolo	3	abstract	OLD
Tu ricordi un evento a Maria	1	abstract	OLD
Maria ti stravolge il pensiero	3	abstract	OLD
Tu provochi una reazione a Maria	1	abstract	OLD
Maria ti esorta allo studio	3	abstract	OLD
Tu riconduci un'azione a Maria	1	abstract	OLD
Maria ti impone una prassi	3	abstract	OLD
Tu progetti una gita con Maria	1	abstract	OLD
Tu scaraventi la boccia a Maria	1	action	NEW
Maria ti getta il tappo	3	action	NEW
Tu passi il gioco a Maria	1	action	NEW
Maria ti impresta il phon	3	action	NEW
Tu dai le carte a Maria	1	action	NEW
Maria ti assegna le mappe	3	action	NEW
Tu affidi le chiavi a Maria	1	action	NEW
Maria ti affibia il cane	3	action	NEW

Tu distribuisci le razioni a Maria	1	action	NEW
Maria ti trasferisce le foto	3	action	NEW
Tu riconsegni l'auto a Maria	1	action	NEW
Maria ti rende la moto	3	action	NEW
Tu sganci lo stipendio a Maria	1	action	NEW
Maria ti sborsa i soldi	3	action	NEW
Tu elargisci denaro a Maria	1	action	NEW
Maria ti fornisce l'ombrello	3	action	NEW
Tu premi con una coppa Maria	1	action	NEW
Maria ti omaggia con un regalo	3	action	NEW
Tu consegni il farmaco a Maria	1	action	NEW
Maria ti somministra la pillola	3	action	NEW
Tu suggerisci un romanzo a Maria	1	abstract	NEW
Maria ti fa alcuni favori	3	abstract	NEW
Tu esponi dei dubbi a Maria	1	abstract	NEW
Maria ti nasconde un segreto	3	abstract	NEW
Tu consigli una vacanza a Maria	1	abstract	NEW
Maria ti protegge gli averi	3	abstract	NEW
Tu descrivi la situazione a Maria	1	abstract	NEW
Maria valuta con te un affare	3	abstract	NEW
Tu insegni l'inglese a Maria	1	abstract	NEW
Maria ti ribadisce un'idea	3	abstract	NEW
Tu autorizzi il trasferimento a Maria	1	abstract	NEW
Maria ti raddoppia il lavoro	3	abstract	NEW
Tu avverti del pericolo Maria	1	abstract	NEW
Maria ti ricorda un evento	3	abstract	NEW
Tu stravolgi il pensiero a Maria	1	abstract	NEW
Maria ti provoca una reazione	3	abstract	NEW
Tu esorti allo studio Maria	1	abstract	NEW
Maria ti riconduce un'azione	3	abstract	NEW
Tu imponi una prassi a Maria	1	abstract	NEW
Maria progetta una gita con te	3	abstract	NEW

F

Old_fixed	Old_moving	new_fixed	new_moving	Phonology Phonemes
MA-DRE	MA-NO			alliterative
NO-CE	NO-DO			alliterative
RA-DIO	RA-NA			alliterative
SA-GRA	SA-LE			alliterative
CO-RDA	CO-RPO			alliterative
VI-NO	VI-TA			alliterative
PA-DRE	PA-STA			alliterative
SE-NO	SE-TA			alliterative
CA-RRO	CA-RTA			alliterative
MA-GO	MA-RE			alliterative
PE-PE	PE-SO			alliterative
SE-GNO	SE-TE			alliterative
TE-LA	TE-STO			alliterative
BA-NCA	BA-NDO			alliterative
RO-BA	RO-GO			alliterative
LI-NO	LI-RA			alliterative
MI-RA	MI-TO			alliterative
LO-DE	LO-GO			alliterative
CI-GNO	CI-MA			alliterative
FA-RO	FA-TA			alliterative
BE-NDA	TI-GRE			dissimilar
FA-LCO	GA-MBA			dissimilar
FI-ORE	TO-PO			dissimilar
AR-MA	LA-TTE			dissimilar
LA-DRO	SP-INA			dissimilar
GA-TTO	SO-LE			dissimilar
ST-UFA	UR-LO			dissimilar
FE-RRO	LU-CE			dissimilar
GI-OCO	RA-TTO			dissimilar
MO-NDO	SC-ARPA			dissimilar
PE-STE	FE-LPA			dissimilar

SF-ERA	LE-TTO			dissimilar
SA-NTO	BA-RCA			dissimilar
ME-SE	FI-ENO			dissimilar
SC-ALA	PR-ATO			dissimilar
GU-FO	FI-UME			dissimilar
VA-SO	TE-NDA			dissimilar
VA-SCA	LA-RDO			dissimilar
TR-ENO	PE-SCE			dissimilar
FU-NGO	CO-STA			dissimilar
		BA-GNO	BA-RA	alliterative
		CE-NA	CE-STO	alliterative
		NA-SO	NA-VE	alliterative
		CA-NE	CA-SA	alliterative
		TO-RRE	TO-RTA	alliterative
		VE-SPA	VE-TRO	alliterative
		MU-RO	MU-SA	alliterative
		MO-DA	MO-TO	alliterative
		SE-ME	SE-RA	alliterative
		NO-IA	NO-ME	alliterative
		AR-IA	CI-BO	dissimilar
		LA-GO	PO-STA	dissimilar
		DI-TO	PI-ZZA	dissimilar
		AS-SO	CA-LZA	dissimilar
		PA-NCA	TA-RLO	dissimilar
		MI-ELE	VI-STA	dissimilar
		PA-RTE	BA-CO	dissimilar
		TE-TTO	PI-EDE	dissimilar
		LI-BRO	GO-NNA	dissimilar
		PE-NNA	CU-ORE	dissimilar