Taking action in hand: Effects of gesture observation on action-verb naming

Ana Murteira

Supervisors: Professor Lyndsey Nickels

Associate Professor Paul Sowman

A thesis submitted in fulfilment of the requirements for the degree of Master of Research

Faculty of Human Sciences

Department of Cognitive Sciences

Macquarie University

April, 2016

Table of contents

List of Tables

List of Figures

Abstract

An extensive body of research suggests a link between action and the mental representation of verbs. Gestures can symbolically represent actions and therefore constitute a bridge between actions and verb processing. This thesis investigates the relationship between meaningful gestures processed in the absence of speech (i.e. pantomimes) and action-verb retrieval. Experiment 1 tested the transparency of a set of gestures in their representation of an action-verb and also investigated possible factors influencing this transparency. Two groups of participants were asked to a) name the action-verb depicted in a pantomime gesture or b) rate the appropriateness of a pantomime gesture in reflecting the meaning of a given verb. Results indicate that, when representing actions, some pantomime gestures clearly map onto a unique action concept, while others are more ambiguous. Pantomime gestures representing actions involving instruments (e.g. *hammer*) are particularly ambiguous. This ambiguity was, however, attenuated when participants had to judge the appropriateness of a pantomime gesture for representing a verb. In Experiment 2, the pantomime gestures from Experiment 1 were used in a cross-modal priming paradigm, to investigate whether the observation of those pantomime gestures had an effect on subsequent action picture naming. Participants were asked to name (using a verb) an action picture, either preceded by a pantomime gesture representing the depicted action (match condition) or by a pantomime gesture representing another unrelated but meaningful action (mismatch condition). Participants were significantly faster at verb naming, in the match condition. Results were also analysed in light of prime-gesture transparency and linguistic factors known to influence verb retrieval (i.e. verb transitivity and instrumentality). The implications of our results in relation to other cross-modal priming effects are discussed and some suggestions made about possible mechanisms underpinning the effects of gesture observation in verb retrieval.

Declaration

I declare that the work in this thesis has not been submitted for a higher degree to any other university or institution. This thesis was prepared by myself, Ana Murteira, under the supervision of Professor Lyndsey Nickels and Professor Paul Sowman. I certify that this thesis is my original work and all sources of information or assistance received have been appropriately acknowledged.

The research presented in this thesis was approved by the Macquarie University Human Research Ethics Committee (Study 1 - Reference: 5201500026; Study 2 - Reference: 5201200035).

Signed:

Ana Murteira

Student number: 44745621

Date: 26th April, 2016

Acknowledgements

I would first like to thank my supervisors Lyndsey Nickels and Paul Sowman. I am grateful to Lyndsey for believing in me, for all patience, generosity, support, guidance and so helpful feedback. I am equally thankful to Paul for clear guidance, feedback, and generosity in sharing knowledge. I am also very grateful to Lesley McKnight for all the “backstage” support and to Andy, for being incredible patience on helping me throughout the project.

I also want to thank to all the amazing people I have the change to meet here: Thank you Oksana, Inga, Hui-Ching and Leonie for your friendship and to make Sydney even more fun!

Thanks to my friends Vania and Adria for encouraging me to pursuit my overseas dream, thank you for all the help and valuable advices. I wouldn’t be here without your support!

I’m also very grateful to all patients and colleagues who I worked with, all these years, as a speech and language therapist. Definitely, the opportunity to work with such an amazing group of professionals, patients and families, motivated even more my curiosity in cognitive sciences.

A special thanks to my family for always be there for me with unconditional love and support. Thanks to my dearest friends that, despite the distance, keep make me laugh! I miss you all very much!

Finally, thank you to my beloved husband, friend and partner for almost twelve years. Thank you for supporting my decisions and share this incredible adventure with me.

Introduction

In recent years, theories about how word concepts are represented and organized in the brain have focused on possible interactions between action and language. A visible way for the body to communicate meaning is through gestures. Hence, understanding the mechanisms underpinning the relationship between gesture and speech can inform theories about how different modalities may contribute to the representation of meaning. While research so far has mostly focused on concomitant gesture and speech production, some empirical evidence demonstrates the existence of strong relationships between meaningful gestures and lexical representations. Nevertheless, little is known about the extent to which this link between gesture and language may influence lexical retrieval: to begin to rectify this gap in knowledge is the aim of our research. Specifically, we aim to examine the interaction between pantomime gesture and verb retrieval.

The introduction section aims to present a critical review of the key developments in gesture and language processing, situating the discussion into the context of current theoretical perspectives. First we will briefly summarize some pertinent theories of conceptual knowledge organization and discuss how they can help us to understand links between action, language and verb representation. Next, we will specifically address the role of gesture in speech and language, highlighting some theories of gesture and speech integration in communication. Finally, we discuss research findings that address the specific type of gesture (i.e. pantomime) that is the focus of this study, and particularly address the recognition of pantomimes.

Action, language and verb representation

Concepts can be described as "mental representations that form the meaning of words and allow us to categorize events and entities in the world" (Medin & Smith, 1984, cited in Bedny, Caramazza, Grossman, Pascual-Leone & Saxe, 2008, p. 11347). For decades, theories have been proposed to explain how concepts are organized and represented in the brain and mind. Some theories hold that the meanings of words are stored in the form of abstract representations that are linked to lexical representations. These abstract representations are organized by conceptual features that are distinct from modality-specific mechanisms for perception and action (see Mahon & Caramazza, 2009 for a review). Another group of theories postulate that conceptual knowledge is grounded in interaction with the world; consequently a concept's representation reflects this interaction and is organized along sensorimotor dimensions (e.g., Barsalou, 2008; Coelho & Bartolo, 2012; Vigliocco, Vinson, Lewis & Garrett, 2004). Such 'embodied' approaches to explaining language processing argue that semantic representations are at least partly constituted by sensory and motor information. Specifically, the neural systems that are involved in forming and retrieving semantic knowledge are argued to overlap with those necessary for perceiving sensory modalities or for producing actions (for review and critique see, for example, Caramazza, Anzellotti, Strnad & Lingnau (2014) and Willems & Casasanto (2011)). An intermediate theoretical position proposes that semantic representations are amodal, and activated in a semantic system that is outside the motor system. However, these representations interact with sensory and motor information in the motor system (Mahon & Caramazza, 2008; Patterson, Nestor & Rogers, 2007).

In their attempt to bring together action and language as part of the same conceptual process, studies taking an embodied perspective give great attention to action-word (action-verb) processing. According to Pulvermüller and colleagues (2005), action-verbs are "abstract semantic links between language elements and motor programs" (p. 977). Verbs require the integration of the event (i.e. what the action is about), the agent (i.e. who is taking the action), the agent's goal and the action's effect, in a coherent representation. These fundamental characteristics cross the boundaries between semantics and syntax. Consequently, theories of verb representation suggest a strong relationship between verb's semantic features and its syntactic properties (Jackendoff, 1990; Rappaport & Levin, 1988). Jackendoff (1990) proposes that verbs are represented by an elementary meaning (likened to a light verb – e.g. *do*) and combined with a number of semantic features to create a more specific meaning. For example, the verb *run* can be represented by the light verb *go* (i.e., specifying the type of event – run is a way of *going*), that is combined with particular semantic features to provide the specific meaning of *run* (Pinker, 1989). These features are associated with the verb's argument structure and the thematic roles of these arguments (that is, what the event is about, who/what is obligatorily involved and in what manner). This grammatical information is known to be activated when a verb is retrieved, not only in a sentence, but also as a single-word (Thompson, Lange, Schneider & Shapiro, 1997). Compared to nouns, relatively little attention has been given to the investigation of semantic representation of verbs. However, research to date has shown that the semantic organization of verbs is less straightforward than that of nouns. In the action domain, semantic features are less correlated within a semantic field, and therefore it might be harder to establish category boundaries (Plant, Webster & Whitworth, 2011; Vigliocco, Vinson, Lewis & Garrett, 2004; Vinson & Vigliocco, 2002, Vinson, Vigliocco, Cappa & Siri, 2003).

Different lines of research have explored the importance of sensorimotor information in action-related language processes (i.e., verbs). If motor regions are involved in the semantic representation of words referring to actions, one should expect fast, relatively automatic engagement of these motor areas in tasks involving activation of the semantic features of these words (Kemmerer, 2015). Indeed, neuroimaging research (e.g., Hauk, Shtyrov, & Pulvermüller, 2004; Kemmerer & Gonzalez-Castilho, 2010; Moseley & Pulvermüller, 2014) and transcranial magnetic stimulation research (e.g., Buccino, Riggio, Melli, Binkofski, Gallese & Rizzolatti, 2005; Innocenti, de Stefani, Sestito & Gentilucci 2014; Oliveri, Finocchiaro, Shapiro, Gangitano, Caramazza & Pascual-Leone, 2004; Pulvermüller, Hauk, Nikulin & Illmoniemi, 2005; Repetto, Colombo, Cipresso & Riva 2013) provide evidence for engagement of motor brain regions during processing of action-language, though the nature of the role played by these areas remains unclear. Nevertheless, such findings have been taken as evidence that (pre)motor regions may be involved in the semantic representation of action verbs (e.g., Buccino et al, 2005; Pulvermüller et al 2005; Repetto et al, 2013). In contrast, other studies find evidence against a specific engagement of the motor system in conceptual and lexical action word processing (e.g., de Zubicaray, Arciuli & McMahon, 2013; Watson, Cardillo, Ianni & Chatterjee, 2013). Between these two extremes, other studies have shown modulation of motor cortical areas only during post conceptual stages of word recognition: that is, explicit retrieval of the motor information associated with action language processing (e.g. a semantic task versus a syllabic task; Papeo, Vallesi, Isaja & Rumiati, 2009) or when a particular strategy is adopted to perform tasks (e.g., participants may imagine the body part movement used to perform the actions described by the verbs; Tomasino, Fink, Sparing, Dafotakis & Weiss 2008; Tomasino & Rumiati, 2013).This suggests that brain motor areas might only be recruited in tasks that require the retrieval of sensorimotor features associated with words, rather than being recruited as an integral part of linguistic processing.

In sum, research to-date does not entirely answer the question of whether motor simulation constitutes an integral part of semantic knowledge or is a secondary effect of cognitive processing related to stimulus and task demands. Some researchers have proposed that, although independent, action and language might interact when sensorimotor information is relevant for specific task resolution. Therefore the involvement of motor brain areas in action word processing is flexible rather than obligatory (Mahon & Caramazza, 2008;

Papagno, Mattavelli, Cattaneo, Romito & Albanese, 2014; Tomasino, Ceschia, Fabbro, & Skrap, 2012; Willems, Özyürek & Hagoort, 2007).

The conceptual and lexical representation of actions has also been studied in language impairment following brain damage (aphasia) (e.g., Kemmerer, Rudrauf, Manzel & Tranel, 2012; Saygin, Wilson, Dronkers & Bates, 2004) and in patients with lesions in the motor system (e.g. Bak et al, 2006; Cotelli et al, 2006; Cotelli et al, 2007; Maieron, Marin, Fabbro & Skrap, 2013). For example, Saygin et al. (2004) examined the comprehension of actions by people with aphasia, in both linguistic domains (i.e., a reading task where patients were asked to complete an action sentence by selecting the appropriate object name) and non-linguistic domains (i.e., a picture matching task, where participants were asked to select the appropriate object picture that matched a picture of an action missing its object). Aphasic participants showed deficits in both domains, but impairments were more pronounced in the linguistic domain with severity of aphasia strongly related to performance in both tasks. Furthermore, when analyzing participant's brain lesions, they found that action interpretation deficits were related to lesions in anterior brain areas known to be involved in motor planning and execution. The authors propose that verbal and non-verbal representations of actions share substrates to a variable degree.

For motor lesions, problems related to the loss of high-level motor representations (e.g., ideational apraxia) would be predicted to have a more severe impact on action lexical-semantic representation than more peripheral motor deficits (e.g. patients with selective deficits in motor areas). The results from studies with patients with lesions in the motor system are far from clear: some researchers have reported evidence for deficits in action naming and action-word comprehension in people with neurodegenerative diseases affecting motor areas (e.g., frontotemporal dementia subtypes: Cotelli et al, 2006; Silveri, Salvigni, Cappa, Della Vedova & Puopolo, 2003; Parkinson´s disease: Cotelli et al, 2007; Herrera & Cuetos, 2012; Herrera, Rodriguez-Ferreiro & Cuetos, 2012; progressive supranuclear palsy:

Bak et al, 2006). However, studies with neurosurgical patients with selective lesions in motor areas have not shown evidence for a direct interaction between this brain region and action-verb naming (Maieron et al, 2013). Moreover, although some of the results from the reported studies could indicate the presence of a specific language dysfunction (i.e. action-verb processing), it is also plausible that tasks implicating verb processing are more cognitively demanding than noun processing and, as a result, greater difficulty in processing verbs is observed (Muslimovic, Post, Speelman & Schmand, 2005; Silveri et al, 2003).

Overall, due to the differences in outcomes across studies, it is not possible to clearly determine the extent to which language and motor systems are dependent or independent systems. Nevertheless, evidence suggests that we can cautiously conclude that language and action are (at least) functionally interactive. This dynamic interplay might be ultimately reflected in the way body movement and, in particular hand posture, can influence language processing (Areshenkoff, Bub & Masson, in preparation; Masson, Bub & Newton-Taylor, 2008; Pine, Reeves, Howlett & Fletcher, 2013; Wheeler & Bergen, 2006). For instance, the results from Pine et al (2013) indicate that producing a hand movement compatible with features of objects facilitates object picture naming (e.g. the production of a flat open-handed movement prior to the appearance of an airplane picture or the production of a closed fist movement prior to the appearance of a hairbrush picture), whereas incompatibility in hand posture slows down object naming. Similarly, Masson et al (2008) have demonstrated that knowledge about gesture representation (i.e. hand posture) is linked to the meaning of words referring to manipulable objects. Hence, we now turn to focus more specifically on the potential for gesture to provide further insights into the relationship between action and language.

Seeking a bridge between action and language: how could gestures help us speak?

Gestures can use body movement to create symbolic meaning, thus providing a bridge between action and language. When gestures represent concrete entities and/or actions through their resemblance to the corresponding events or objects, they are considered to be iconic (McNeill, 2006). Iconicity has recently become more of a focus under an embodied view of language, in which iconicity works as the key to linking linguistic forms and conceptual representations based on perception and motor experiences (Perniss, Thompson & Vigliocco, 2010; Perniss & Vigliocco, 2014; Vinson, Thompson, Skinner & Vigliocco, 2015). In spoken language, iconic gestures have been investigated in co-occurrence with speech, that is, co-speech gestures. Co-speech iconic gestures use body movement, particularly movement of the hands, while a person speaks, often used to express information that may not be simultaneously conveyed verbally, such as object size or characteristics (e.g. a speaker talking about a cake and produces a gesture for a round shape) or event/object motion (e.g, rolling down event described in Kita & Özyürek (2003)). However, iconic gestures can also occur in the absence of speech, where they are usually referred to as pantomimes. In pantomimed gesturing, body parts engage in a pattern of action with common features related to the actual pattern of action (e.g., a person miming an action such as drinking)[1].Emblems are another type of gesture frequently used in the absence of speech: these are highly conventionalized gestures that are immediately associated with an idea or word (e.g., thumb up recognized as OK; McNeill, 2000; Scharp, Tompkins, & Iverson, 2007).

Cognitive psychology and linguistics have attempted to explain how gesture and spoken language are linked during communication. While some authors argue that gesture and speech form separate systems (e.g. de Ruiter, 2000; Kita & Ozyürek, 2003; Krauss, Chen & Gottesmann, 2000) and others suggest they are part of a single system (Hostetter & Alibali,

[1] The example refers to a context where a pantomime gesture is produced in the absence of speech. However, the presence of speech is optional (please refer to Kendon (2004) for examples of pantomime gestures in co-occurrence with speech production).

2008; McNeill, 2000), both agree that gesture and speech form an integrated system – they are semantically and pragmatically co-expressive (Kita, 2000; Kita & Özyürek, 2003); gestures can be temporally organized with speech (Kelly, Özyürek & Maris, 2010; Morrel-Samuels & Krauss, 1992); they share a tight relationship that is reflected in the hand which gesture is produced (more often the right hand), which might suggest a link with the left hemispheric speech system (Corballis, 2003); gestures have an effect on how speech is perceived (Beattie & Shovelton, 1999; McNeill, Cassell & McCullough, 1994); and gesture and speech neurally overlap in brain areas involved in processing of both (for a review see Özyürek, 2014).

There are two main theoretical positions that are based on the assumption that gesture and speech are separate systems: one suggests that gestures arise from conceptual encoding and have a complementary communicative role to spoken language (e.g., Sketch Model; de Ruiter, 2000); the other suggests that gestures do not carry meaning on their own, but derive it during lexical retrieval (e.g. Lexical Retrieval Hypothesis; Krauss et al, 2000). As noted above, other accounts share the assumption that gesture and speech are parts of one system, because they arise from a common underlying cognitive system (Goldin-Meadow, 2003; Hostetler & Alibali, 2008; McNeill, 1992). The next section attempts to provide an overview of three of these theoretical models and illustrate the differences between them.

Models of gesture-speech production – separate systems perspective

The Lexical Retrieval Hypothesis (Krauss et al, 2000) and Sketch Model (de Ruiter, 2000) have incorporated mechanisms for gesture production into Levelt's speech production model (Levelt, 1989; Levelt, Roelofs & Meyer, 1999) and share the assumption that speech and gesture are separate, but interactive systems (see Figure 1). Four subsystems are involved in processing gestures and language– a conceptual level, responsible for generating the message that a speaker wishes to communicate; a gesture planner, responsible for generating the gesture; a formulator, responsible for language processing and a level responsible for the activation of visuo-spatial images. The main difference between the models relates to the mechanisms by which, and levels at which, gesture is considered to interact with speech production.

In the Sketch model (Figure 1a), gesture initiates in the conceptualizer. A gesture sketch is formulated at this level. This sketch conveys visuospatial information retrieved from working memory and/or retrieving a gesture template (i.e., abstract knowledge about gesture shapes) from a store of gestural conventions (*gestuary)*. This sketch extracts the necessary information for planning the gesture. At the same time, a pre-verbal message is constructed at the same conceptual level and sent to the formulator for language production. The sketch gesture is sent to the gesture planner, which generates a motor program. This allows the motor program to be adapted according to contextual constraints (for example, adapt the gesture to perform it with one rather than two hands when one hand is occupied). The gesture planner feeds back to the conceptual level to coordinate gesture and speech timing (see de Ruiter, 2000, for further specific details). According to this model, although independent systems, gesture and speech interact at the level of conceptualization and, therefore, they are part of the same communicative intention. This model does not deny that gestures could facilitate speaking process; however such effect is likely to occur only indirectly (via conceptualization)

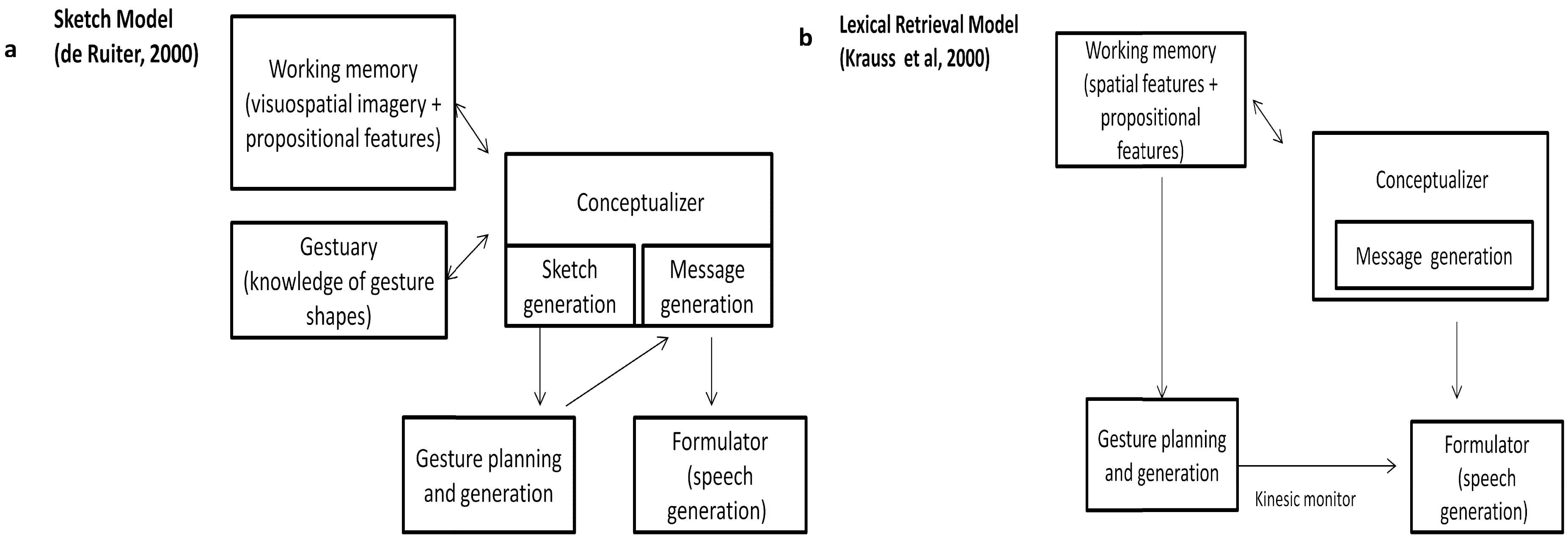

*Figure 1***.** Simplified representations of two models of language and gesture production: a) Sketch model (adapted from de Ruiter, 2000); b) the Lexical Retrieval Hypothesis model (adapted from Krauss et al, 2000).

In contrast, the Lexical Retrieval Hypothesis (Krauss et al, 2000; Figure 1b) argues for a direct role of gestures in the process of lexical retrieval. Iconic gestures, which the authors call “lexical gestures”, are generated from spatial features from the speaker’s working memory. These abstract features are the input to a motor planner, which translates them into a set of instructions about how to execute the gesture. The gesture is activated at the moment of speaking, via cross-modal priming between the gesture’s spatial features and lexical - conceptual features (see Krauss et al, 2000 for further specific details). According to this model, gestures “are a product of the same conceptual process that ultimately results in speech, but the two systems diverge early on” (Krauss et al, 2000, p. 272), and therefore they do not share the same conceptualizing stage, when the preverbal message is constructed.

Model of gesture-speech production – same system perspective

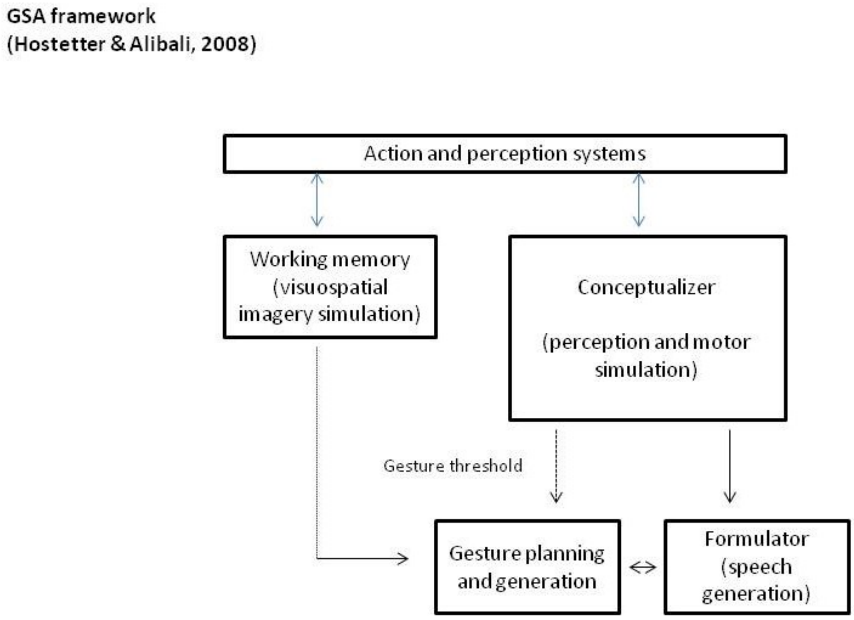

Figure 2. Gesture as Simulated Action framework (adapted from Hostetter & Alibali, 2008).

Based on the embodied approach, a different view about the possible mechanism that gives rise to gestures is proposed by the Gesture as Simulated Action framework (Figure 2; Hostetter & Alibali, 2008). This framework proposes that gestures and speech are part of the same cognitive and communicative system, because both are expressions of the same simulation process: “speaking involves simulations of perception and action; forming a simulation evokes a motor plan [initiated in working memory] that come to be expressed in gesture” (Hostetter & Alibali, 2008, p. 509). Gesture production is dependent on three factors: 1) the strength of activation of the simulated action, which must be strong enough to spread from motor areas and exceed the gesture threshold; 2) the current speaker’s gesture threshold (i.e. while speaking, some individuals rely more on gestural strategies than others) and 3) the engagement of the motor system while speaking. The GSA framework focuses on explaining how gestures arise from speech and accepts as possible all the other models’ proposals, by which gesture might benefit speech.

Gestures without speech: The case of pantomime gesture

The models above are based on gesture production or comprehension in co-occurrence with language, as is most of the work supporting the different positions regarding the relationship between iconic gestures and language. This is important because, as noted by Goldin-Meadow & Wagner (2005), gesture changes in function and in form when produced on its own compared to when it is performed as a complement to language. For instance, under the circumstance where speakers use gesture alone to communicate, gesture assumes more language-like characteristics, abandoning the more abstract format that characterizes gesture when co-occurs with speech (Goldin-Meadow, 2007).

Pantomime gestures involve hand movements representing objects or actions (Rose, 2006) and occur and can be understood without accompanying speech (Willems, Özyürek & Hagoort, 2009). Pantomime gestures are interesting in the sense that they might have a different relationship to the verbal message they attempt to represent compared to co-speech gestures or emblems. Unlike emblems, they do not seem to obey linguistic constraints or conventions (McNeill, 2000), however, they must be sufficiently informative to provide meaning independently. In contrast to co-speech iconic gestures, it is well accepted that pantomimes are intended to be communicative and are communicatively effective (Krauss et al, 2000). For representing an object or an action through a pantomime, people have to rely on their mental representation of that object or event, selecting conceptual features that meet the constraints of gesture, that is, features that represent physical or spatial properties of the concept (Nispen, Sandt-Koenderman, Mol & Krahmer, 2014). The research on production of pantomimes depicting objects or tool-use actions has demonstrated that different speakers tend to use similar gestures for representing the same objects or actions (Masson-Carro, Goudbeek, & Krahmer, 2015; Nispen et al, 2014; O'Reilly, 1995). In gesture recognition, people are often able to recognize the general meaning of the action represented in a pantomime, although some specifics can only be disambiguated with additional contextual information. Osiurak and colleagues (Osiurak, Jarry, Baltenneck, Boudin & Le Gaill, 2012) asked participants to identify the tools used to perform actions from pantomimes. Participants correctly identified the tool more often in a multiple choice task, than when required to directly name it from the gesture. Osiurak et al. (2012) suggest that recognizing objects from pantomimes without additional contextual information is difficult, because different individuals may have different representations about what features of the gesture must be critically present.

Cognitive models of gesture comprehension and production in the absence of co-occurring speech have been developed in the field of limb apraxia. As an example, we present the model of praxis processing proposed by Rothi, Ochipa and Heilman (1997; Figure 3). Unlike the gesture-speech production models, this model does not attempt to explain the possible integration of gesture and speech. The fundamental aim is to explain the mechanisms of skilled gesture processing, from comprehension to production, based on studies of apraxic participants. Nevertheless, we propose that this model can be readily adapted to account for the close relationship between gesture and speech.

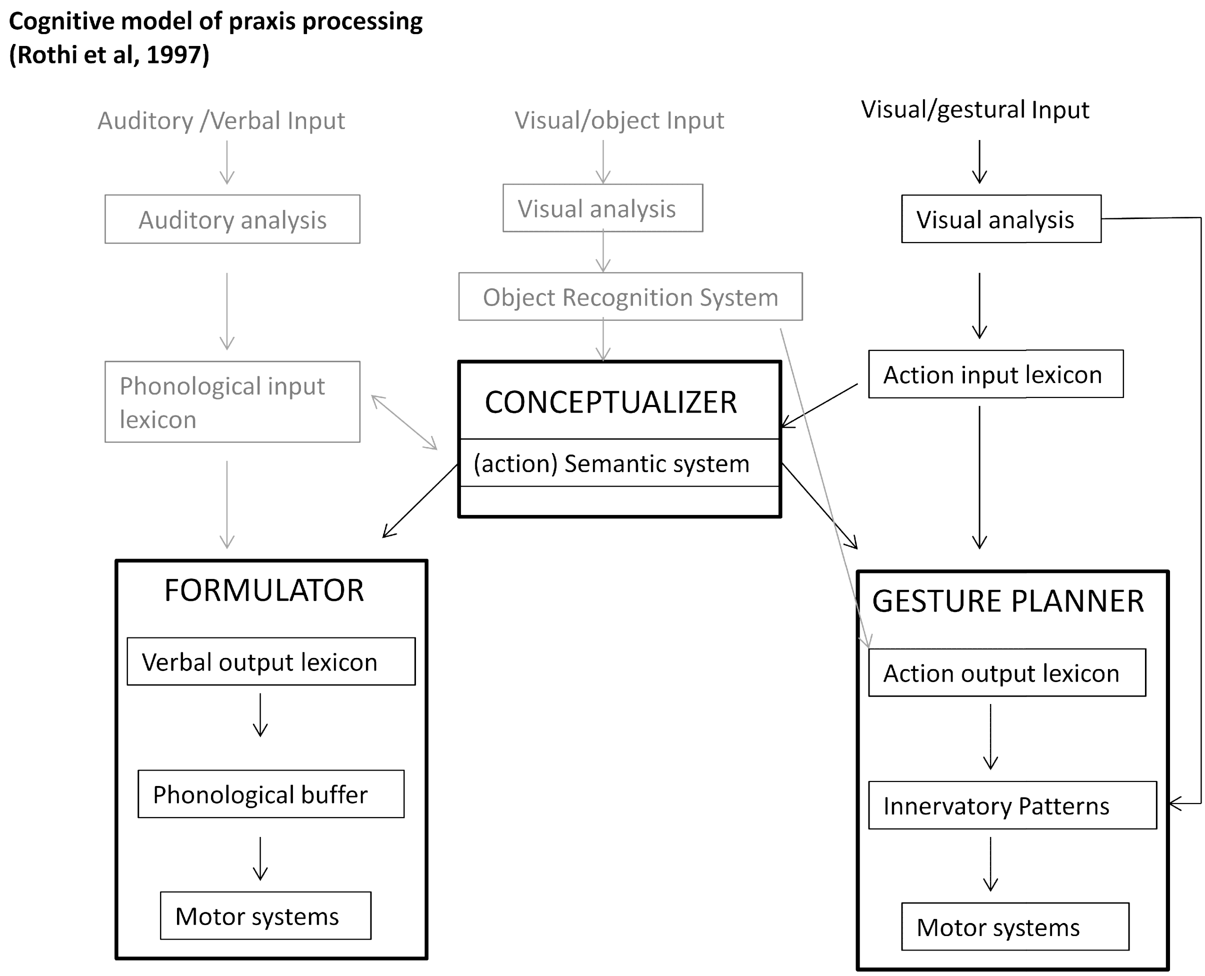

Figure 3. A model of praxis processing and its relation to semantic, naming and word and object recognition (adapted from Rothi et al, 1997). The figure was adapted to integrate the levels of speech and gesture production based on the models previously described

According to this model, when a gesture is perceived, it activates information about the physical attributes of the gesture, which is stored in the action input lexicon. If the perceived gesture is recognized as familiar, there will be an activation of the knowledge about that action, which occurs at the level of semantics. If a familiar gesture is to be produced, activation flows from the action semantic system to the action output lexicon. The output lexicon contains the representations of the physical attributes of a to-be-performed gesture. These representations are subsequently transformed into motor programs for execution.

At the centre of the model is the action semantic system, a conceptual component of the praxis system that involves knowledge of the functions of tools and objects, knowledge of actions and knowledge about the organization of actions sequentially (Rothi et al, 1997). The action semantic system is, in the authors view, a subsystem within the broader conceptual-semantic system. The model accounts for a link between the gestural and language system through semantics. A shared conceptual system between gesture and language is also assumed by some of the gesture-speech production models described above (e.g. de Ruiter, 2000). Rose (2006) proposes that de Ruiter's *gestuary* may play the same role as the "action output lexicon" in Rothi et al's model, because both involve the stored knowledge for the production of emblems and emblem-like pantomimes (e.g. use a hammer). According to de Ruiter (2000) the conceptualizer accesses the *gestuary* in the search for any conventional gesture available to represent a concept. It is also possible to reason that this knowledge of gesture shapes is not only available for production, but also for gesture recognition. Therefore, the action input lexicon, in Rothi's et al model may also overlap with the proposed concept of *gestuary*, but for gesture recognition.

Effects of pantomime on spoken language

While the models of de Ruiter (2000), Krauss and colleagues (2000) and Hostetter & Alibali (2008) focus on gesture that occurs together with speech, Rothi et al's (1997) model focuses on gesture production in the absence of speech. In recent years, there has been interest in the link between iconic gestures and language, even when gestures are presented alone. The available body of research suggests that there is indeed an effect of iconic gesture production or perception on language processing (Bernardis & Caramelli, 2009; Bernardis, Salilas & Caramelli, 2008;; Cohen-Maximov, Avirame; Flöel & Lavidor, 2015; de Marco, Stefani & Gentilucci, 2015; Yap, So, Yap, Tan & Teoh, 2011; Mounoud, Duscherer, Moy & Perraudin, 2007). For example, Bernardis and colleagues (2008) asked participants to watch a pantomime video-clip and then read a word (noun or verb) as fast as possible. They report significant inhibition when the word to read was preceded by an unrelated pantomime, however no facilitation was found from the related condition. They conclude that the delay in language processing, when word and gesture were mismatched, indicates that gesture and language interact. Following the unexpected lack of facilitation in the matching condition, Bernadis and Caramelli (2009) conducted another experiment using the same materials and conditions as Bernardis et al (2008) but, this time, participants had to form a mental image of the concept evoked by the words before responding: This manipulation resulted in facilitation from pantomime primes in the matching condition. Bernadis and Caramelli (2009) concluded that in this condition pantomimes and words activated the same visual-spatial and perceptual information, thus producing a repetition priming effect. Yap and colleagues (2011) further explored the gesture-word priming effect, finding that iconic gestures prime lexical decision to semantically related words. In contrast to Bernardis and Caramelli (2009), who suggested a repetition priming effect due to shared visuo-spatial and perceptual information, Yap et al. (2011) explain their results in light of the semantic spreading activation framework (Collins & Loftus, 1975, cited in Yap et al, 2011). Additionally, So and colleagues (So, Yi-Feng, Yap, Kheng, & Yap, 2013) explored differences in the influence of iconic gestures alone (Gesture-only condition) and co-occurring with speech (Gestures-accompanying-speech condition) on word recognition (i.e., lexical decision). They report that only the Gesture-only condition lead to significant priming effect. The authors suggest that processing iconic gestures alone might result in deeper semantic processing: As iconic gestures do not have strong conventional labels, they might be connected to a wide range of semantic concepts, strengthening the priming effect in lexical decision.

To our knowledge, only two studies have addressed the influence of pantomime perception on picture naming. These studies examined object (tool) recognition in adults (Moy & Mounoud, 2003) and children (Mounoud et al, 2007). In both studies, participants were found to be faster to name a tool when previously primed by a pantomime of someone pretending to manipulate that tool, suggesting a relationship between tool recognition and action representation.

In summary, it is clear that action and language are closely linked. However, there remains debate as to the nature of these links at both neural and cognitive levels. Many alternative theories have been proposed regarding the nature of the relationship between iconic gestures and language (e.g., single versus independent systems) and the role of gesture in communication. These theoretical accounts differ in their nature and extent of interaction between the systems. Moreover, it is also recognised that action represented in gestures has a close and important relationship to language, but once again there is no decisive theory regarding how the two may interact. The evidence is particularly lacking regarding the influence on language processing of gestures produced without language. Pantomime gesture's lack of conventionality, on one hand, and ability to convey meaning without a mandatory verbal context, on the other hand provide an interesting perspective on the study of the links between gestural and spoken language, particularly in the case of action verbs. Hence, in this study we aim to provide further empirical data to allow further specification of these relationships.

The present study

The aim of the present study was to further explore the relationship between pantomime gesture and verb retrieval across two experiments.

In the first experiment we examined how well people named and recognized actions depicted in gestures. Studies addressing pantomime naming and recognition have mainly come from pantomime tasks in the field of apraxia (e.g., Negri, Rumiati, Zadini, Ukmar, Mahon & Caramazza, 2007; Rothi, Heilman & Watson, 1985) and aphasia (e.g., Ferro, Santos & Caldas, 1980; Wang & Goodglass, 1992). Less attention to this issue has been given in healthy participants, and the sparse literature has only addressed recognition of pantomimes depicting object use (e.g, Buxbaum, Kyle & Menon, 2005; Osiurak et al, 2012). Hence, in

Experiment 1, we extended investigation of gesture naming and recognition to include different types of actions beyond just object-use actions, and examine factors that influence gesture naming accuracy. This experiment also allowed us to select gestures for the second experiment.

Experiment 2 constitutes the principal experiment of this study and aimed to analyze the interplay between pantomime gesture and verb retrieval. Specifically, we investigated whether pantomime gestures would prime verb retrieval, in a naming task. Previous research has shown that iconic gestures when presented alone or paired with speech can prime semantically related words in lexical decision tasks (Yap et al, 2011; So et al, 2013) or facilitate object recognition (Mounoud et al, 2007). This experiment provides new evidence for cross-modal priming of verb retrieval in picture naming and, for the first time, explores the effect of pantomimes with different levels of transparency on lexical retrieval.

Experiment 1

The aim of this experiment was to establish the transparency of selected pantomime gestures and investigate the potential linguistic factors that may influence this transparency. The pantomimes in this experiment were also to be used as primes in Experiment 2.

Two measures were used to determine how well a gesture represented an action (verb): gesture naming accuracy and a gesture-name appropriateness rating. Using these two measures it was possible to explore the differences between: (a) how far it was possible to derive an action concept from a gesture directly – how far it evoked a specific action-verb and, (b) how far a gesture was consistent with an action concept evoked by a specific verb. In addition, we were interested in exploring which factors contributed to naming accuracy. We reasoned that there might be an influence on gesture naming of the conceptual and lexicosyntactic properties of the to-be named action-verb. In the past decade, there have been several investigations of the influence of conceptual and grammatical factors on verb retrieval, including instrumentality, transitivity, and frequency (e.g., De Bleser & Kauschke, 2003; Jonkers, 1998; Jonkers & Bastiaanse, 2007; Kauschke & von Frankenberg, 2008; Luzzati et al, 2002; Thompson et al, 2007). Given the possible interface between gesture and action-verb retrieval, it was our goal to explore if these factors also influence gesture naming.

Method

Participants

Participants were recruited using Amazon's Mechanical Turk platform cohort to participate in two online surveys. The study was approved by the Macquarie University Human Research Ethics Committee and all participants provided informed consent.

Experiment 1A - Gesture naming: Thirty native English speakers (16 male; see Table 1), participated in this experiment. All participants had completed at least High school education.

Experiment 1B – Gesture-name appropriateness rating: Thirty native English speakers (17 male; see Table 1), who had completed at least High school education, participated in this experiment; however one participant was excluded due to non-valid responses (that participant provided the same rating for every item in the task).

Table 1.

Sample average age

	Age group (frequency %)			
Task	18-25	26-45	46-59	≥ 60
1A Gesture naming	13%	77%	6.7%	3.3%
1B Gesture-name rating	10%	72%	10%	6.9%

Materials

Surveys were compiled using Qualtrics software (Qualtrics, Provo, UT). Gesture video clips were created and used as stimuli. Each video consisted of a person miming an action (gesturing). Some of the gestures represented actions performed with objects, without the physical presence of the object (e.g. *combing*), whereas other gestures represented actions that were not object-oriented (e.g., *walking*; see Figure 4 for example stills from the video clips). Gestures were chosen that represented actions in the action pictures that were to be used as target stimuli in Experiment 2. Transparent gestures (i.e., those where non-signers may be able to infer the meaning of a sign) from sign language databases (British Sign Language: bslsignbank.ucl.ac.uk/dictionary and Australian Sign Language: auslan.org.au) were used to assist in gesture development. All gestures were videoed for this project using the same actor. Then these potential pantomimes were tested for recognition, by presenting the clips to a small sample of participants. Of the 86 pantomime gestures, 11 were poorly recognised by 8 of the 10 participants and were amended. All gesture videos were edited into clips of 900ms duration using Windows Movie Maker software.

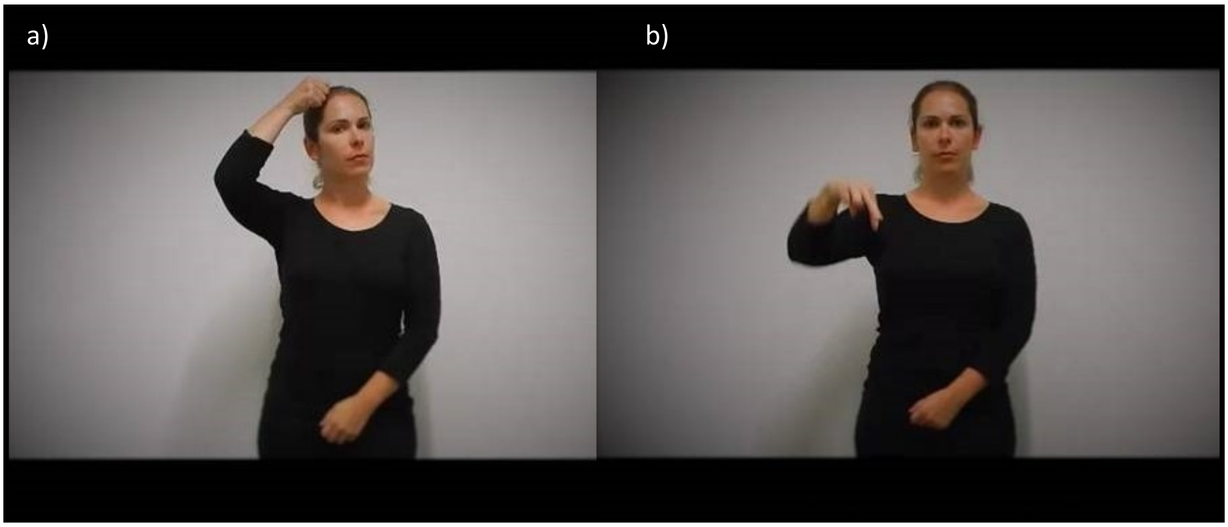

*Figure 4***.** Examples of stills from the video clips of pantomime gestures used in the task; a) combing; b) walking

Procedure

Experiment 1A - Gesture naming: A pantomime gesture video clip was presented in the centre of the screen with a single line for response entry below the video. Participants viewed the gestures (black screen for 500ms followed by 900ms of the target gesture) and entered (via keyboard) a single verb that best described the action depicted by each gesture. They did this for each of the 86 pantomime gestures in turn. Opportunities to rest were provided during the task.

Experiment 1B - Gesture-name appropriateness rating: This task required participants to rate how well pantomime gestures reflected the meaning of an action-verb. Participants viewed 86 pantomime gestures as short video clips (black screen for 500ms followed by 900ms of the target gesture). For each gesture, a single verb was presented on the screen, simultaneous with and below the video. After watching the pantomime gesture video and reading the verb, participants were asked to indicate their level of agreement or disagreement with the question "How well does the depicted gesture reflect the meaning of [verb]". Participants were required to indicate their response using the following scale: 1- not at all; 2-very little; 3-quite well; 4-very well; 5- perfectly. Participants could only continue to the next item after they had selected a value on the scale. Seven catch trials, conditions where gesture and verb totally mismatched, were randomly introduced throughout the trial sequence. These trials were used to judge if participants were paying attention to the task.

Results

Experiment 1A - Gesture naming task

Only target names were accepted as correct. Synonyms were considered incorrect responses, as they were likely to have different values for psycholinguistic variables of interest (e.g., word frequency). Mean accuracy of gesture naming was calculated for each item. To analyse response error types, naming errors[2] were classified as (based on Rumiati & Humphreys (1998)): (1) *visual errors* – naming an action which had a visual movement similar to that of the target action, but was not from the same semantic category (e.g., *write → scrape)*; (2) *semantic errors* – naming an action which was from the same verb semantic category or represented the same concept (e.g. *whisper → talk*); (3) other – naming an action visually and semantically unrelated to the target (e.g. *pay → shuffle*) or non verb responses (e.g. *swing → side to side*).

Figure 5 represents the mean gesture-naming accuracy and mean proportion of each type of naming error and Figure 6 illustrates the differences in naming accuracy for each gesture, ordered by accuracy (Appendix A provides details on accuracy and error types for each item). Most gestures were named with greater than 80% accuracy (n=33).Semantic and visual errors occurred with approximately equal frequency (16% and 18% respectively).

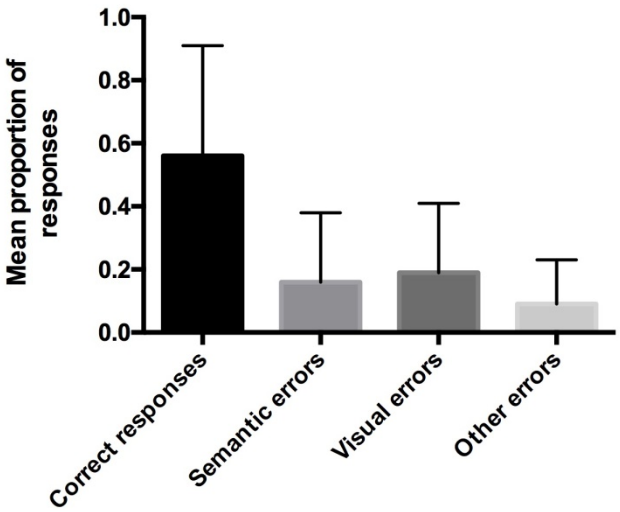

Figure 5. Proportion of correct responses and error types in the gesture naming task

[2] We used the term 'error' to refer to a non-target response, and appreciate that some responses (and synonyms in particular) could be considered acceptable alternatives rather than erroneous responses. Synonyms only constitute a small proportion of participant's responses, and reanalysis including synonyms as correct responses only slightly to .60 and mean sightly decreases proportion of semantic errors to .12. Reanalysis using this coding shows the same pattern as those reported using a strict coding.

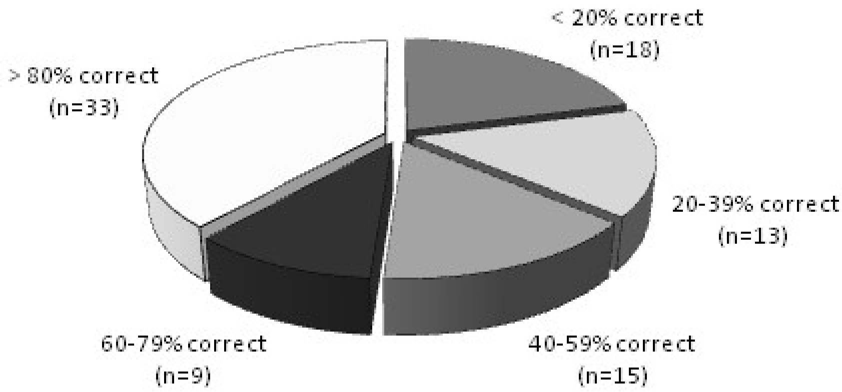

Figure 6. Distribution of gesture naming accuracy

Experiment 1B - Gesture-name appropriateness rating task

The overall mean name appropriateness rating for all gestures was 3.72 (*SD*=.85) with the majority (47.7%) of items receiving a mean rating of between 4 ('very well') and 5 ('perfectly'; see Figure 7).

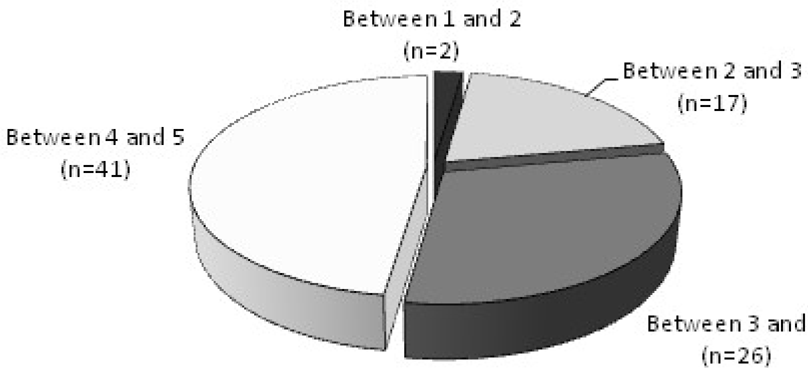

*Figure 7***.** Mean gesture-name appropriateness rating distribution across the rating categories. The categories represented in the figure correspond to each rate level of how well a pantomime gesture reflected the meaning of an action-verb: 1– not at all; 2-very little; 3- quite well; 4-very well; 5- perfectly

Relationship between gesture naming accuracy and gesture-name appropriateness rating

The relationship between how well a gesture was rated as representing a given verb and accuracy on naming the same gestures was investigated using Pearson's correlation coefficient (see Figure 8). Overall, there was a strong positive relationship between the name rating and gesture naming accuracy (r=.805, n=86, p<0.001): Higher ratings were associated with an increase in naming accuracy.

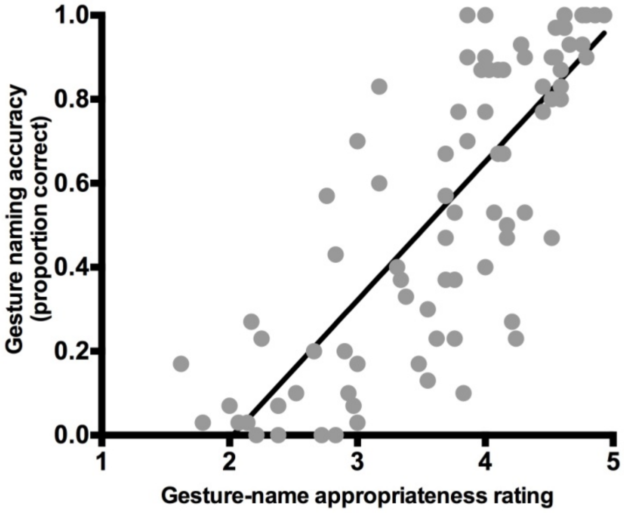

Figure 8. Relationship between gesture naming accuracy and gesture-name appropriateness rating

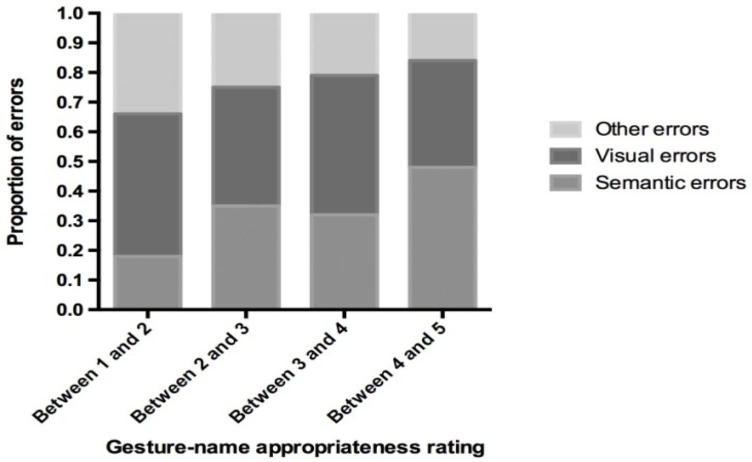

Figure 9. Prevalence of error types within each gesture-name appropriateness rating category; the categories correspond to each rate level of how well a pantomime gesture reflected the meaning of an action-verb: 1– not at all; 2-very little; 3- quite well; 4-very well; 5- perfectly

A closer look to the distribution of naming errors (Figure 9) showed a prevalence of visual errors in all levels of gesture-name appropriateness rating, with exception of gestures judged, on average, in category "*between 4 and 5"*. This category represents those gestures with a high level of gesture-name appropriateness rating and high naming accuracy, and although the proportion of naming errors was small (only 19% of total responses constituted naming errors), there was a prevalence of semantic errors.

Influence of psycholinguistic variables on gesture naming accuracy

We explored whether some of the factors known to influence verb retrieval also influenced the variation in gesture naming accuracy: Multiple regression was performed to investigate the relationship between gesture naming accuracy and psycholinguistic variables of word frequency, transitivity and instrumentality (see Table 2).

Table 2.

Intercorrelations (zero-order Pearson's r) for naming accuracy and psycholinguistic variables and regression analysis summary examining the effects of frequency, transitivity and instrumentality on gesture naming accuracy.

	M	*SD*	Freq. r	Trans. r	Instr. r	*B*	*SE B*	*ß*	Part *r*
Naming accuracy	.56	.35	-.126	-.279**	-.303**				
Word frequency (Log)	1.46	.60	--	.111	.047	-.055	.060	-0.95	-.095
Transitivity	.69	.47		--	.365**	-.137	.083	-.184	-.170
Instrumentality	.37	.48			--	-.165	.079	-.231*	-.215

** $p < .001$; *p < .05

The correlation between the independent variables was sufficiently low as to suggest an absence of multicollinearity (Field, 2005). Using Tabachnick & Fidell's (2013) guidelines for inspection of Mahalanobis distance, no outliers were found. Zero-order correlations show that there was a small to moderate, negative correlation between transitivity and gesture naming accuracy, suggesting that overall, gestures representing intransitive actions were better named (M = .70, *SD* = .33) than gestures representing transitive actions (M =.49, *SD* = .34). There was also a moderate, negative correlation between instrumentality and gesture naming accuracy, suggesting that gestures representing non-instrumental actions were better named

(M=.64, *SD* = .35) than gestures representing instrumental actions (M =.42, *SD* =.31). There was no significant correlation between word frequency and gesture naming accuracy.

In the regression analysis, transitivity, instrumentality and word frequency explained, as a whole, 13.3% of the variance in gesture naming accuracy, (F (3, 82) = 4.2, R^2 = .133, p = .008). When comparing the independent contribution of each variable (i.e., comparing the independent contributions of transitivity and instrumentality), only instrumentality significantly predicted the variation in naming accuracy. Figure 10 summarizes the results for the rate of correct responses and naming errors across gestures representing instrumental and non-instrumental verbs. As a significant correlation between transitivity and instrumentality was found, only gestures representing transitive verbs were included in this analysis: 29 gestures representing non-instrumental verbs and 30 gestures representing instrumental verbs.

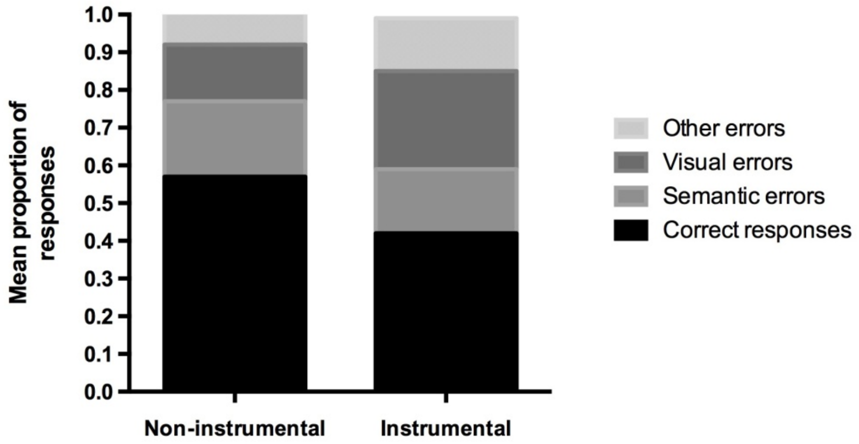

Figure 10. Proportion of correct responses and naming errors for gestures representing instrumental and non-instrumental verbs

The analysis of the proportion of semantic and visual errors indicates that participants produced a significantly higher proportion of visual errors when naming gestures representing instrumental verbs (t (721.8) = 6.24, p < .01; Instrumental: M = .44, *SD* = .23, Non-instrumental; M= .34, *SD* = .29) and, correspondingly, a significantly higher proportion of semantic errors when naming gestures representing non-instrumental verbs (t (744.8) = 7.33, $p < .01$; Instrumental: M= .29, *SD*=.29; Non-instrumental = .46, *SD*= .36).

Discussion

The first goal of Experiment 1 was to explore how well a set of gestures represented a corresponding set of actions. To achieve this goal, participants were asked to name the possible action depicted by a gesture or rate how well a gesture represented the meaning of a given action verb. Unsurprisingly, there was variability of naming accuracy across gestures, with participants producing both visual and semantic errors. Nevertheless, in the gesture-name appropriateness rating task, most of the gestures were rated within the categories of *"very well"* or *"perfectly"* regarding how well they represented a given verb. Moreover, there was a positive relationship between naming accuracy and name-appropriateness with gestures rated as more representative of a given action-verb meaning also eliciting higher naming accuracy. However, for some gestures, participants presented with poor naming, but gave a high rating in the gesture-name appropriateness task. A qualitative analysis of the naming errors produced by these gestures indicates a number of semantic naming errors, characterised by the production of close semantic neighbours to the target verb (e.g. *yell – shout/talk*; *sew – knitting*) or near synonyms (e.g., *watch - see*).

On the one hand, pantomime gestures are thought to be reflective of communicative intent in the sense that they can express semantic information about the event they aim to represent, without requiring a co-occurring verbal expression (Scharp et al, 2007). Indeed, the results of the naming task showed that people can identify pantomime gestures as meaningful and recall lexical-semantic information about the action enacted. However, on the other hand, pantomimes are also considered idiosyncratic (McNeill, 2000): Because pantomimes do not have a totally conventionalized form, different individuals can represent pantomimes in different ways. Thus, if the mental representation of a pantomime is different across subjects, their ability to retrieve the specific meaning, when observing someone depicting that gesture, can also differ.

In the present experiment, variation of naming accuracy across different gestures presents evidence for some degree of gestural idiosyncrasy, in the extent to which those gestures are clearly informative about the specific action verb they represent. Nevertheless, the effects of gesture idiosyncrasy seem to be attenuated when participants only have to judge to what extent a gesture appropriately reflects the meaning of a given verb. This result is in agreement with the context hypothesis (Osiurak et al, 2012), which suggests that performance in pantomime recognition is better with contextual information. In the case of the present experiment, the contextual support was given when participants were provided with the specific verb enacted by the gesture. However, some gestures have sufficiently clear and unambiguous mappings to action concepts that no further context was required.

Gesture transparency reflects the relationship between gesture features and the representation of the target action-verb. Linking a gesture to an action-verb requires activation by the gesture of those conceptual-semantic features of the action that can be physically represented with body-movement. Therefore, some features, important to specify the meaning of the to-be-named verb, cannot be fully represented through gesture (e.g. a boxing ring or boxing gloves to distinguish between *boxing* and *fighting*). In this situation, gestures are still recognized as meaningful actions, but may be insufficiently specified to uniquely map onto a verb and, consequently, a naming error is more likely to occur. Errors in gesture recognition have been investigated in normal aging (Ska & Croisile, 1998), showing a prevalence of semantic errors in older adults when naming gestures. In our experiment, the overall numbers of visual and semantic errors were very similar.

We further explored the factors underlying variation in gesture naming accuracy, and specifically whether it was influenced by properties known to influence verb retrieval: the grammatical property, transitivity; the lexical property, word frequency; and the conceptual property, instrumentality. We found no effect of word frequency and, when exploring the independent contribution of transitivity and instrumentality, only instrumentality significantly predict gesture-naming accuracy. Participants' naming was better for gestures representing non-instrumental action-verbs, than for gestures representing instrumental action-verbs. Instrumental verbs are actions which require an instrument (not being a body part) and it is assumed that, for these verbs, the instrument is part of the conceptual representation of the verb (Jackendoff, 1990, cited in Jonkers & Bastiaanse, 2007). When eliciting a verb depicted by a pantomime, if is not clear what instrument is used to perform the action, it might be harder to retrieve the specific verb. In our results, visual errors were the most frequent type of naming error for gestures representing instrumental actions. Therefore, the difficulties in naming a gesture representing a specific verb are likely to reflect a conceptual error, because a similar body movement could represent actions executed with different tools (e.g. *typing* on a keyboard versus *playing* piano).

In sum, the results from Experiment 1 indicate that pantomimes, although communicative, form a heterogeneous group. Some are clearly transparent in regard to the action represented, whilst others are more context-dependent. Gestures representing instrumental actions seem particularly susceptible to the need for contextual support to be fully recognized. Our results are in agreement with the suggestion that gestures representing instrumental actions are more ambiguous (Osiurak et al, 2012).

Experiment 2

The results of Experiment 1 suggest that some pantomime gestures clearly convey the semantic features of the actions that they are intended to represent. In contrast, other pantomime gestures provide a less clear-cut referent and consequently may be lexicalized as different verbs. These findings, therefore, provide evidence for different levels of transparency in action representation by gesture. The present experiment seeks to further explore the link between pantomime gestures and verbs. It does so by examining the extent to which the observation of pantomime gestures has an effect on action picture naming.

Method

Participants

Thirty-six healthy participants (13 males and 23 females; age range 18-48 (M=27.78, *SD*= 8.53) were recruited from the participant databases of the Departments of Cognitive Science and Psychology at Macquarie University. Participants were provided with either monetary compensation or course credit. All participants were English native speakers with normal or corrected-to-normal vision and no reported history of neurological or psychiatric illness. The study was approved by the Macquarie University Human Research Ethics Committee.

Materials

This experiment involved presentation of gesture primes followed by action pictures for elicitation of target verbs.

Verb Targets: Targets were verbs corresponding to pictures of actions retrieved from the International Picture Naming Project (Szekely et al., 2004), the Object and Action Naming Battery (Druks & Masterson, 2000) and a new version of the Verb and Sentence Test (Bastiaanse, unpublished). All images were 300x300 pixel, black-and-white line drawings. From a total set of 395 pictures of actions, 89 pictures were selected for which it was possible to produce a pantomime gesture, and for which the original source reported name agreement above 70%. Of these 89 pictures, 86 had Australian name agreement (de Aguiar, 2015). The final set comprised those 80 pictures of actions with Australian name agreement above 80%.

Gesture Primes: 80 gesture video clips, generated as detailed in Experiment 1, were used as primes in the experiment. The remaining six video clips from Experiment 1 were used as practice trials.

Prime-target pairs: Target action picture stimuli were presented in two conditions: prime-target match or prime-target mismatch. In the match condition, the gesture (prime) expressed the same action as the verb (target) depicted in the action picture; in the mismatch condition, the gesture expressed a different action to that of the action picture. The related primes and targets were recombined to create the mismatch pairs ensuring that the action represented by the gesture (prime) and the target verb were not associated (using the Edinburgh Associative Thesaurus; EAT, Kiss, Armstrong, Milroy and Piper, 1973, available online at http://www.eat.rl.ac.uk) nor from the same verb category (using VerbNet; Kipper-Schuler, 2006). Additionally, in the mismatch condition, prime-target pairs did not start with the same phoneme and prime-target pairs that were visually similar in terms of action body movement were avoided.

Two lists were generated, each of which presented half of the targets in the match condition and half in the mismatch condition. The lists were matched for different psycholinguistic properties of the targets (i.e, word frequency, age of acquisition, imageability, transitivity, regularity, number of phonemes, instrumentality and initial phoneme). Each participant was randomly assigned to a list, and saw each target only once (e.g., participant 1 would be presented with target picture X in the prime-target match condition and participant 2 would be presented with target picture X in the prime-target mismatch condition).

Procedure

The experimental task was programmed using Presentation® software (Version 16.3, www.neurobs.com). Gesture video clips and pictures were presented serially, in the centre of the computer monitor, against a black background. Each trial comprised: (a) a fixation point (+) appearing in the centre of the screen for 500ms, (b) a prime video clip for 900ms; (c) a central fixation point (+) for 200 ms; (d) a target picture displayed for 2000 ms (see Figure 11). Participants were asked to pay attention to the video clip and respond, as fast and accurately as possible, to the second stimulus (i.e., the action picture), naming each item with a single verb that represented the action in the picture. Each participant performed six practice trials and 80 experimental trials (40 target stimuli in the prime-target match condition and 40 target stimuli in the prime-target mismatch condition). The experiment lasted a maximum of 12 minutes. The order of stimulus presentation was randomized for each participant. Vocal responses were recorded and reaction time was measured from picture display onset. Participants' responses were saved as sound files and verified for timing and accuracy manually using Audacity® software 2.1.1 (available at http://www.audacityteam.org/) to ensure accurate vocal reaction time measurement.

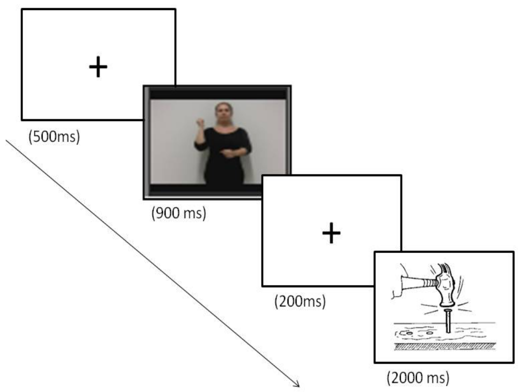

Figure 11. Trial presentation structure.

Data analysis

Eight pantomime gestures from Experiment 1 with both naming accuracy below 20% and a mean gesture-name appropriateness rating below 3 were excluded from analysis in this experiment. The target *combing* was also excluded, as the majority of participants named this action as *brushing,* which was a target response for another picture in the task. Therefore, from a total of 80 possible pairs, 71 pairs remained for analysis. Removing these nine items did not alter the psycholinguistic matching across sets. Only correct target verbs were accepted. Potentially acceptable alternative naming responses (e.g. name *seeing* instead of *watching*) were scored as incorrect and excluded from further analysis as such responses would have different psycholinguistic properties to the target (e.g. word frequency, number of phonemes). In addition, trials with items where reaction times had a z-score above ±3.29 were considered outliers and excluded from analyses (Tabachnick & Fidell, 2013). Missing cases, inaccurate responses and outliers together constituted 10.2% and 8.7% of the cases excluded from the mismatch and match conditions, respectively.

The dependent variable examined was response latency (as measured by reaction time from picture onset) and the independent variable was the prime type (prime-target match and prime-target mismatch). Paired-samples t-tests were conducted to evaluate the effect of pantomime gesture primes on verb naming latency both by-items and by-participants.

Results

Priming effects

Table 3 presents the mean and standard deviations of reaction times and accuracy in the two conditions at the item and participant levels. Accuracy was almost at ceiling in both conditions and was not analysed further.

Table 3.

Reaction time (ms) and accuracy (proportion correct) of verb production in action picture naming when primed by matching or mismatching gestures.

	Reaction Time				Accuracy			
	Match		Mismatch		Match		Mismatch	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Item level	861	135	1005	124.6	.92	.26	.90	.30
Participant level	873	127.4	1015	124.6	.92	.05	.90	.06

At the item-level, there was a significant difference in mean latency (t (70) = 8.657, p < .001 (two-tailed)) with the mismatch condition being 144 ms slower on average than the match condition.

Although the initial design was a between-subject design, that is, each participant named an item of interest in only one condition (e.g., participant 1 named item X only in the matched condition, whereas participant 2 name item X only in the mismatched condition), given that we matched the targets in the two conditions within a list, we used a within-subject design for the by-participants analysis. Once again the difference between conditions was significant, t (35) = 11.17, $p < .001$ (two-tailed), with an overall advantage of 142 ms for the match condition.

Differences between match and mismatch conditions were also analysed for each item, to explore, in more specific detail, which gesture-verb pairs showed greater priming. A homogeneity test showed that there was a significant difference in the extent of priming across verbs (H=187.4, $p < .001$): For 49 of the 71 items (70%), the mean response latency was significantly faster in the match condition than in the mismatch condition (see Appendix B). However, 19 items showed no significant facilitation (e.g. *pushing*, match 840 ms, mismatch 873 ms), including 11 items which produced slower responses in the match condition of which three showed significant inhibition (e.g. *sliding*, match 885 ms, mismatch 770 ms).

Priming and gesture iconicity/transparency

We further explored the relationship between the amount of priming and the iconicity and/or transparency of the gesture. A partial correlation was used to examine the relationship between gesture naming accuracy from Experiment 1 (i.e. name agreement of the gesture) for the match gesture prime and the size of the priming effect (difference between verb naming latency between the match and mismatch conditions), while controlling for name agreement of the mismatch gesture prime.

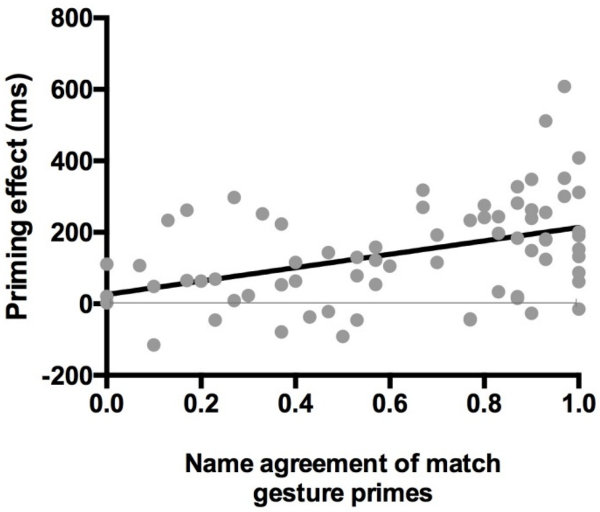

Figure 12. Relationship between size of the priming effect (verb naming latency for the prime-target mismatch condition minus verb naming latency for the match condition) and name agreement of the match gesture prime while controlling for name agreement of the mismatch prime

There was a medium, positive and significant correlation between the amount of priming in Experiment 2 and name agreement of the match gesture prime controlling for name agreement of the mismatch gesture prime (r= .41, n=68, p< .001; see Figure 12): Verbs which showed larger priming effects had correspondent gestures that were significantly more accurately named in Experiment 1. Name agreement of the mismatch gesture did not significantly correlate with priming (although there was a negative trend, r= -.12, n=68, p=.14).

We also analysed the influence of type of gesture naming error (semantic or visual) on the priming effect.

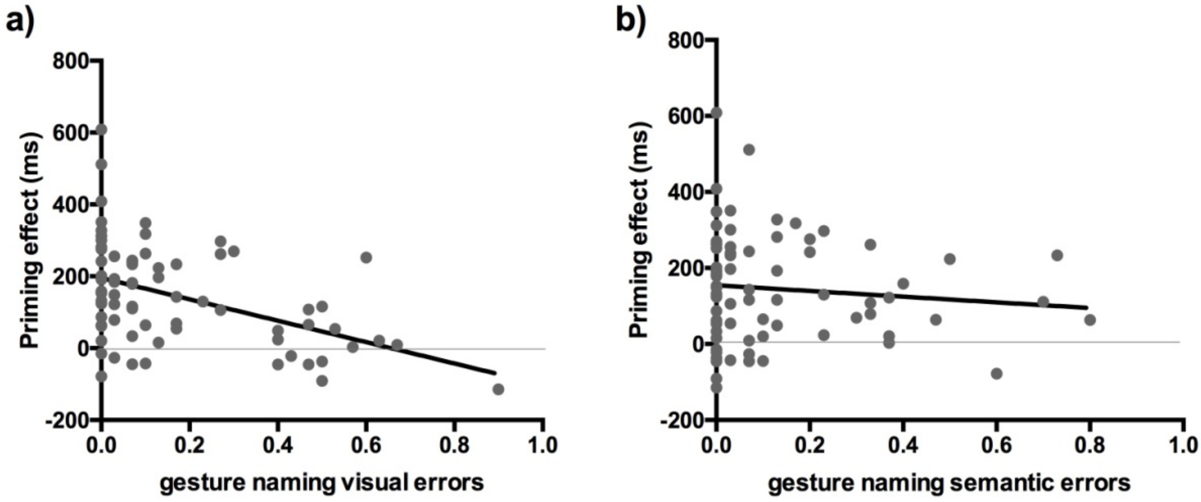

Figure 13. Relationship between size of the priming effect and a) proportion of visual errors in gesture naming (from Experiment 1) and b) proportion of semantic errors in gesture naming (from Experiment 1)

There was a significant, medium sized, negative correlation between the proportion of visual errors to a match gesture prime and the size of the priming effect, r=-.45, n=71, $p <$ .001 (see Figure 13): The higher the proportion of visual errors the smaller the priming effect. For semantic errors there was a small trend in the same direction but this was not significant, r=- .10, n=71, p= .40.

Effects of verb transitivity and instrumentality

We next explored whether the priming effect was influenced by the grammatical (transitivity) and conceptual (instrumentality) properties of the verb. A mixed ANOVA was computed including the factors: Condition (match, mismatch) and Psycholinguistic variable (transitivity: transitive/intransitive; instrumentality: instrumental/non instrumental). As these two variables are correlated, each analysis controlled for the other variable by using subsets of the stimuli: To evaluate the effect of transitivity 22 intransitive and 23 transitive verbs were selected, all of them non-instrumental; To evaluate the effect of instrumentality, 23 instrumental and 23 non-instrumental were selected, all of them transitive.

a) Transitivity Analysis

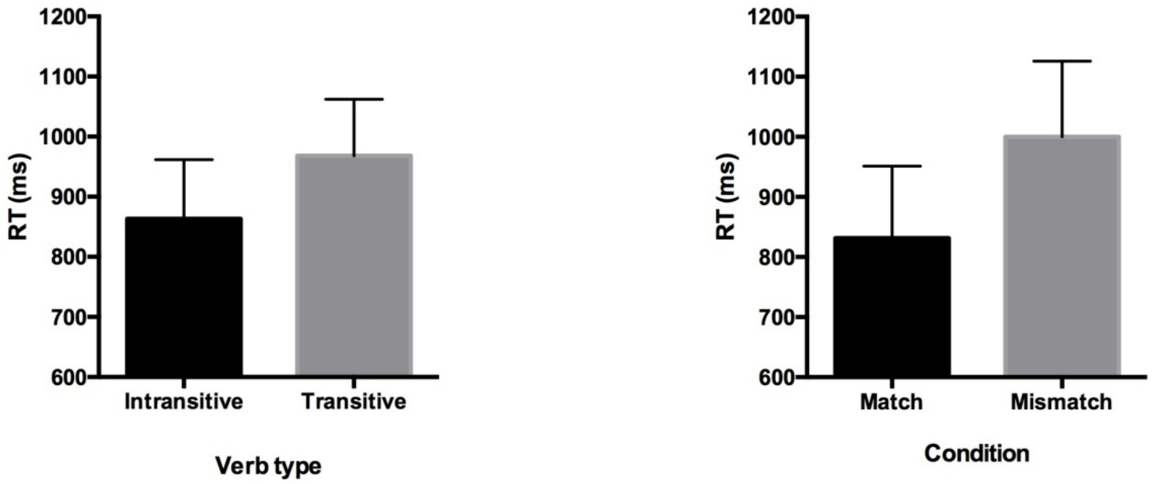

Figure 14. Main effects of Verb type (transitivity) and Condition for transitive and intransitive verb naming latency. Error bars = Standard Error

As shown in Figure 14, there was a significant main effect of Condition (match vs mismatch), F (1, 43) = 54.6, *p* < .001, partial eta squared = .56, with an overall mean difference of 169 ms between naming latency in the match and mismatch conditions. The main effect of transitivity was also significant, F (1, 43) = 13.3, *p* = .001, partial eta squared = .24, with an overall mean difference of 105 ms between transitive and intransitive verb naming latency. There was no significant interaction between Transitivity and Condition, F (1, 43) =.96, *p* = .332, partial eta squared = .02. Post-hoc t-tests confirmed that the difference in reaction time between transitive and intransitive verbs was significant within the match condition (t (43) = - 3.55, *p* = .001 (two-tailed) and within the mismatch condition (t (43) = - 2.19, *p* = .034 (two-tailed)). Intransitive verbs were named faster (M = .77, *SD* = .12) than transitive verbs (M= .90, *SD* = .12) within the match condition, as well as within the mismatch condition (intransitive: M = .96, *SD* = .14; transitive: M = .10, *SD* = .11).

b) Instrumentality Analysis

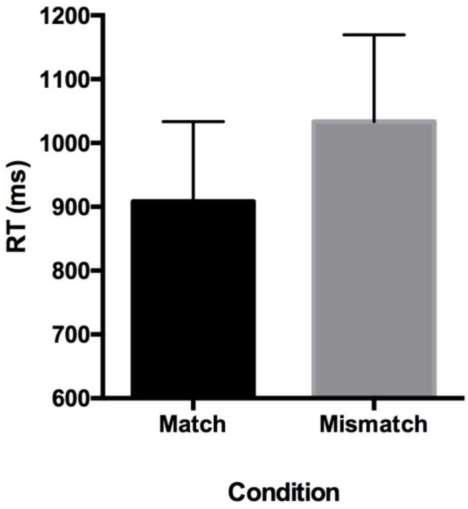

Figure 15. Main effect of Condition for instrumental and non-instrumental verb naming latency. Error bars = Standard Error

As shown in Figure 15, there was a substantial main effect of Condition (match vs mismatch), F (1, 44) = 51.9, $p < .001$, partial eta squared = .54, with an overall mean difference of 125 ms between naming latency in the match and mismatch conditions. However, there was no significant main effect of instrumentality, F (1, 44) = 0.3, $p = .864$, partial eta squared = .001, nor an interaction between Instrumentality and Condition, F (1, 44) = 1.50, $p = .228$, partial eta squared = .03[3].

[3] Results from Experiment 1 showed that Instrumentality predicted gesture naming accuracy. In addition, in Experiment 2 the amount of priming was influenced by gesture naming accuracy. Therefore, we also ran the Instrumentality analyses controlling for gesture naming accuracy, to explore: a) if there was an interaction between condition and gesture naming accuracy - there was not (Condition*Gesture naming accuracy, F (1,43) = 2.1, p=.15) and b) if controlling for this variable affected the interaction between condition and instrumentality - it did not – this interaction remained non-significant (Condition*Instrumentality, F(1,43) = 0.6, p=.45).

Discussion

The main goal of Experiment 2 was to explore the effect of pantomime gesture on action naming and, specifically, to determine whether the observation of pantomime gestures primed action picture naming. Participants were asked to name a depicted action with a verb either when preceded by a gesture representing the same action or when preceded by a gesture representing an unrelated meaningful action. Results revealed faster latencies at action picture naming when primed by a gesture representing the to-be-named action, compared to when primed by a different meaningful gesture. This finding is consistent with other studies that have shown cross-modal gesture-word priming for object naming (Bernardis et al, 2009; Mounoud et al, 2007) or in lexical decision for nouns (e.g., Yap et al, 2011; So et al, 2012). However, this study is the first to present evidence for cross-modal gesture-word priming of verb production.

In this study we used a measure of priming that is defined as the difference between match and mismatch conditions (Neely, 1991). However, as noted by Posner & Snyder (1975, cited in McNamara, 2005) using this comparison poses a problem, because one cannot determine whether there is facilitation in the match condition, inhibition in the mismatch condition or a combination of both. A solution to this problem is to use a 'neutral' baseline condition. Unfortunately, in this experiment, it was not possible to find a neutral prime condition, matched closely to the primes but that would not interfere with processing of verb naming (Jonides & Mack, 1984). For example, even meaningless gestures could comprise potentially meaningful elements that may influence with subsequent processing in verb naming.. Because we are primarily interested in the overall priming effect, that is, whether pantomime gestures influence verb retrieval, the inclusion of a neutral prime condition is therefore not critical (Jonides & Mack, 1984). In addition, our results showed that the naming accuracy of the mismatch gestures (i.e., those ones used as unrelated primes) did not significantly correlate with the size of the priming effect. Therefore, it seems unlikely that the observed priming effect is driven by an inhibition effect from the mismatch condition, but favours that there is facilitation in the match condition.

Some models of gesture-speech production assume that iconic gestures affect lexical retrieval (e.g. Krauss et al, 2000); however, these models are intended only to account for gestures that occur with concomitant speech. Moreover, Krauss and colleagues (2000) go even further, arguing that a gesture must be performed for lexical facilitation to occur. Findings from the present study, along with others (e.g. Mounoud et al, 2007, Marangolo et al, 2010), suggest that the processing of a semantically congruent gesture is enough to enhance lexical retrieval: No gesture production is required.

Our results can be explained within any theory of feature-based semantic representation (e.g., Collins & Loftus, 1975; McRae, de Sa & Seidenberg, 1997; Plaut, 1995). It seems plausible that properties observed in the gesture activate semantic features of a particular verb. Consequently, when the to-be named action picture is displayed, that semantic representation is already partially activated, producing priming. In contrast, when gesture and picture represent different actions, as in the mismatch condition, the semantic features activated by the gesture are unrelated to the target, hence no priming occurs. However, it is also possible that the priming occurs not from the overlap in semantic features but instead from the gesture prime activating the lexical representation of the target verb. This would be consistent with the mechanism proposed for identity priming in spoken word production (e.g. Barry, Hirsh, Johnston & Williams, 2001; Wheeldon & Monsell, 1992).

In contrast to previous studies (Mounoud et al, 2007; Bernardis et al, 2008), we used gesture primes with different levels of transparency (as reflected in different levels of name agreement). By doing this, it was possible to examine to what extent priming effects were influenced by the accessibility of the verb from the gesture, which may enable us to distinguish between the two potential priming mechanisms above. Krauss and colleagues (2000) suggest that not all gestures facilitate verbal production equally well, arguing that only gestures that derive from semantic features that are part of the lexical item will facilitate lexical retrieval. Our findings support Krauss and colleagues' statement: The more transparent the gesture prime was, and hence the more consistently it evoked the target verb, the greater the priming effect. We also found that gestures that, when named, produced a high rate of visual errors, were associated with less priming. In contrast, there was no relationship between proportion of semantic errors elicited in gesture naming and the size of the priming effect. This seems somewhat counterintuitive as it is well established that semantically related word primes result in inhibition of picture naming (Howard, Nickels, Coltheart & Cole-Virtue, 2006; Wheeldon & Monsell, 1994), and gestures which tend to evoke semantic errors could therefore be thought of as analogous.

Why might gestures that evoke semantic errors fail to result in semantic inhibition? First, we should note that semantic inhibition in word production has only been established for nouns, and the different semantic properties of verbs may mean that the effects are less straightforward. For example, some authors have argued that, compared to nouns, for words referring to events, more features are shared between members of diverse semantic fields (e.g. verbs of different categories, such as *walk, drive, roll, sweep,* share the feature <involves motion>) and, therefore it is harder to establish clear boundaries between fields (Vinson & Vigliocco, 2002). An additional contributing factor to the lack of inhibition may be that gestures can be ambiguous. The features of an event that can be physically represented through gesture may not be sufficient to distinguish between the target verb and several other verbs semantically (or visually) related to the target verb. This will result in co-activation of many items with similar semantic features, including the target word, rather than a single highly activated semantic competitor (as in the standard semantic inhibition paradigm).[4] In other words, many items may be primed at the lexical level including the target. Therefore, when the picture is presented and activates the target there is no net inhibitory effect. For

[4] Nevertheless, the constraints of the gesture naming task are such that participants will produce a verb when required to do so.

example, the gesture representing *boxing* produced a high rate of semantic errors in naming, particularly *fighting* or *punching*. When this gesture was presented as prime, for the picture depicting the same action (i.e. *boxing*), features representing fighting, punching and boxing may be activated and the corresponding lexical representations equally activated.

Importantly, in addition to semantic inhibition from semantic competitors at a lexical level, a number of authors have suggested that there is short-lived semantic facilitation from overlap of semantic features (e.g. Vitkovitch, Rutter & Read, 2001; Wheeldon and Monsell, 1994). This mechanism would also be consistent with our results. The overlap in semantic features between gestures (even ambiguous gestures) and verb targets may be enough to lead facilitation of verb naming.

We have demonstrated that gesture transparency is an important factor in facilitating action verb naming. However, it could also be the case that syntactic and conceptual characteristics of the verb play a role. It has been argued that lexical-semantic representations for verbs contain information about their argument structure and instrumentality (e.g., Jackendoff, 1990). Transitivity is related to the verb argument structure information that is activated at the lemma level (Thompson et al, 1997; Jonkers, 1998). Our results show that the observation of a related gesture facilitated naming actions for both transitive verbs (e.g. *tying*) and intransitive verbs (e.g. *walking*). In addition, intransitive verbs were always named faster than transitive verbs. These results are in line with other studies that show that transitive verbs are generally more difficult to retrieve in action naming (De Bleser & Kauschke, 2003; Thompson et al, 1997). However, there was no interaction between transitivity and priming - the cross-modal priming of verb retrieval by gesture is not affected by this verb property.

When considering verb instrumentality, we found no indication of an influence of instrumentality on either verb naming latency or the effect of gesture priming. What is interesting, however, is that, in Experiment 1, instrumentality predicted gesture naming accuracy: We found higher gesture naming accuracy for gestures representing non-instrumental verbs (e.g. *swimming*) than for gestures representing instrumental verbs (e.g. *hammering)*. We reasoned that, because of the constraints imposed by the gestural modality, some instruments could not be clearly represented and, as a consequence, these actions induced more gesture naming errors. As gestures representing instrumental verbs seem to be more ambiguous, one might have expected differences in priming between instrumental and non-instrumental verbs. However, both were primed equally by the related gesture.

In conclusion, the results from Experiment 2 provide evidence that observing gestures influences action word production. We suggest that this cross-modal priming effect is consistent with both priming at a semantic level and/or at the lexical level (see below further discussion). Our results are also in line with previous evidence focusing on nouns suggesting that semantic information contained in congruent hand postures can enhance word processing (e.g. Masson, Bub & Newton-Taylor, 2008) or can facilitate object naming (Pine et al, 2012).

General Discussion

The aim of this research was to provide a better understanding of the mechanisms underpinning the interface between pantomime gesture and language production. Specifically, we aimed to examine the interaction between gesture and verb retrieval. A first experiment investigated gesture naming and gesture recognition. It further examined some linguistic and conceptual factors that might influence gesture naming accuracy. The second experiment investigated the effect of pantomime gesture observation on lexical retrieval, by analysing the extent to which pantomime gestures primed verb retrieval in a picture naming task.

Experiment 1 showed that although pantomime gestures are indeed communicative, they are not always fully comprehensible without a verbal context. Consequently gesture naming is susceptible to naming error. The visual and movement features encoded in gestures encompass the semantic content that these gestures aim to represent. However, when the movement does not convey sufficient of the semantic features required to identify a specific action (verb), errors occur in naming. This was particularly evident for actions where instruments are key to discriminating between visually similar actions (e.g. *combing* or *brushing*; *typing* or *playing the piano*), but are hard to represent in the gesture. Nevertheless, the semantic features that are conveyed by such gestures are still consistent with their verb targets and hence name-gesture appropriateness ratings are high. The findings of this experiment are in agreement with other authors' claims of an iconicity continuum (e.g., Feyereisen, Van de Wiele & Dubois, 1998): Some gestures have an explicit meaning, some gestures are too rudimentary to be interpreted, and those in between would vary in the clarity of their meaning.

In Experiment 2, using a cross-modal priming paradigm, we found that observing pantomime gestures of an action can facilitate the lexical retrieval of verbs expressing that action. However, the degree of facilitation was dependent on the transparency of the gesture, that is, the extent to which a particular gesture is uniquely nameable as the target lexical item.

As stated in the introduction, the theories that have been proposed to explain the relationship between gesture and speech are focused on simultaneous gesture and speech production in the same speaker (e.g. de Ruiter, 2000; Kita, 2000; Krauss et al, 2000). Because they only consider gestures produced at the same time as speaking, they have very little to say about how gesturing in the absence of speech may impact speech production. Nevertheless, these models provide valuable suggestions as to how gesture and speech might be integrated: either by sharing the same conceptual level (e.g. de Ruiter, 2000) or by facilitating lexical retrieval via cross-modal priming at the word form level (Krauss et al, 2000). In the remaining discussion, we propose some mechanisms by which naming of pantomime gesture could occur and by which observation of gestures could facilitate action naming.

According to Rothi's cognitive model of praxis processing, gesture and language processing systems interact at the conceptual-semantic level. The information at this level is thought to be prelinguistic, and represents the meaning of our concepts (Bierwisch & Schreuder, 1992, cited in Nickels, 2001). For some authors (e.g. Dell, Schwartz, Martin, Saffran & Gagnon, 1997; Plaut, 1995: McRae, de Sa & Seidenberg, 1997) the conceptual-semantic system is comprised of a set of semantic features that may include sensory and motor information (Barsalou, 2008; Coelho & Bartolo, 2012; Vigliocco et al, 2004). When familiar gestures are observed, they activate semantic features that are part of the conceptual representation of an action. The subsequent processing steps leading to the naming of a gesture are the same as those used to retrieve a word (e.g. Levelt, 1999; Levelt et al, 1999).

As we suggested in the Discussion of Experiment 2, gestures could facilitate action-verb naming due to priming (persisting activation) of overlapping semantic features or due to priming of the lexical form. However, the fact that we found differences in the size of the priming effect among target verbs means that gestures influence action naming to different extents. How could this be implemented theoretically? How could a gesture activate a verb's semantic and/or lexical form? Drawing on the theoretical accounts in the literature, we suggest three possible mechanisms: a) a stored gesture template activates conceptual-semantic features; b) a stored gesture template directly activates a lexical representation; and c) the visual and movement features of a gesture activate corresponding conceptual-semantic features.

In the Sketch model of gesture-speech production, de Ruiter (2000; see Figure 1, earlier) proposes that, to produce pantomimes, speakers must access stored knowledge about how to produce that gesture. Rose (2006) establishes parallels between de Ruiter's proposal and the action output lexicon in Rothi and colleagues' model of gesture processing (1997; see Figure 3, earlier), as both represent a store for production of conventionalized and emblem-like pantomime gesture movements. Rothi et al. also propose an input store for the form of familiar gestures which is activated when these gestures are observed. We suggest that, for the more transparent gestures, the processing of the visual and movement features leads to activation of a stored gesture template in this input store[5]. This gesture template activates a specific action conceptual representation (represented as a series of features) which in turn maps onto a single lexical word (Figure 16 – pathway 1). Alternatively, the activated gesture template could directly activate a corresponding lexical item (Figure 16 – pathway 2). This could be an amodal lexical representation such as a lemma (e.g., Levelt, et al, 1999), in the same way that such representations are activated by spoken or written input. For more ambiguous gestures, the gestural visual and movement features of that gesture may activate several gesture templates and consequently conceptual-semantic features consistent with a number of concepts will be activated. These, in turn, will activate several possible lexical candidates, including the target word.

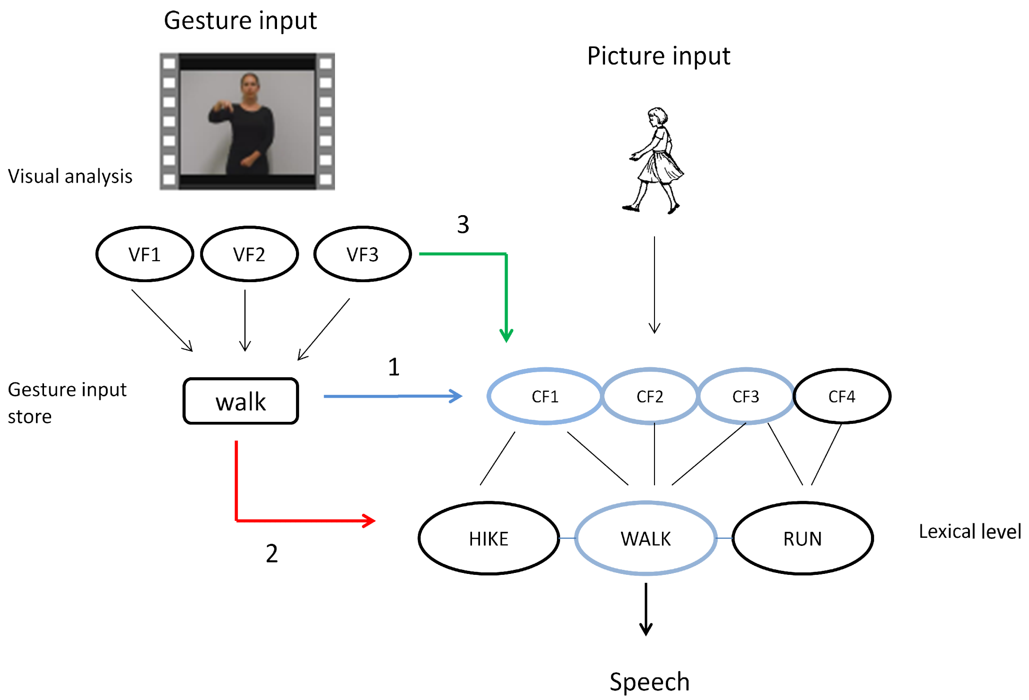

Figure 16. Illustration of the possible mechanisms underlying facilitation of verb production from pantomimed gestures a) from a stored gesture template to conceptual-semantic features (pathway 1, blue arrow); b) directly from a stored gesture template to lexical forms (pathway 2, red arrow); c) directly from visual features to

[5] Even if gestures do not have a single specific stored form, common to all speakers, that does not mean that different speakers might not have different stored forms for the same gesture: This would also result in variability.

conceptual-semantic features (pathway 3, green arrow). [6]VF = gesture visual features; CF = conceptual-semantic features

The accounts above assume that pantomime gestures have a stored form. However, in contrast to Rothi et al. (1997), McNeill (2000) argued that pantomime gestures do not have a conventionalized form, although some may be more emblematic than others. In other words, they do not have a stored form. Findings from our experiment, along with others (e.g., Osiurak et al, 2012) have shown that people are far from perfect when naming pantomimes. Therefore, although people can attribute a meaningful label to the gesture, expressing knowledge about what it could represent, they may not have that gesture representation as a single unit. For these reasons, we suggest that it is also possible that conceptual-semantic features could be directly activated by the visual and movement features of gestures (Figure 16 – pathway 3). This also seems plausible given that the same gestural movement can be used for a wide variety of actions (e.g. similar arm motion can be used to represent different actions, such as knocking on a door, ringing a handbell, painting a wall). This proposition receives some support from studies showing a link between spatial and motoric representations and semantic processing (e.g., Masson et al, 2008; Areshenkoff et al, 2015; Dudschig, de Filippis & Kaup, 2014). Nevertheless, although our findings are consistent with visual features of gesture directly activating conceptual-semantic features, it also is clear that there needs to be sufficient overlap of features to yield a priming effect.

The mechanisms here proposed might reconcile the different positions in the literature regarding a conceptual or lexical role for gesture production in speech production. Krauss and colleagues (2000) propose a direct route from gesture to a lexical level. In their model, gestures (as a set of motoric features) can directly facilitate speech production, as the motor features encode part of the lexical item's meaning (e.g. *round* as a feature of ball). This idea is

6 The framework represents the likely stages of processing in the cross-modal priming task from seeing a gesture to seeing a picture, and hence appears to be primarily top down. However, it is not our intention to make any claims about the interaction between levels and modalities of processing. For example, the links between semantic and lexical levels could either be unidirectional (e.g., Levelt et al, 1999) or interactive (e.g., Dell et al, 1997).

contested by de Ruiter (2000) who argues that facilitation from direct lexical activation is only possible if the gestural features specify a single lexical item. He suggests that if the gesture contains features associated with more than one lexical item, incorporating a mechanism by which there is direct activation of the lexical level, would slow down selection of the target item. Therefore, de Ruiter argues that gestures can only facilitate speech indirectly, via shared conceptual processing. We believe that the two views may not be mutually exclusive as our data can account for both proposals, which seems dependent on the transparency of the observed gesture.

More generally, our findings can be accounted for by assuming shared underlying processing between gesture and language related to action. According to Hostetter and Alibali (2008), gestures are the product of internal action simulation activated during speech. Hostetter and Alibali models' only accounts for co-speech gesture, however pantomime gestures may also reflect action simulation even in the absence of speech. A simulation "is a recreation of the neural states involved in performing or witnessing a particular action, in the absence of actually doing so" (Hostetter & Alibali, 2010, p. 245). Thus, gestures, as meaningful hand actions, mirror the real actions and their iconic aspect may activate conceptual features that have an impact on subsequent language processing. In a recent meta-analysis of fMRI studies, Yang and colleagues (2015) suggest that comprehension of meaningful gestures involves the same perceptual-motor network important for action recognition, mapping the observed gestures onto the motor representation in the brain. In the same way, an extensive body of research has shown that perception and action systems are involved in the semantic representation of verbs (for review see: Meteyard et al, 2012). Additionally, meaningful gestures convey semantic information and, therefore their comprehension is associated with brain areas involving semantic processing and integration (Willems et al, 2009; Yang et al, 2015), in common with language processing.

In light of our results, we propose that gestures and action-verbs converge at the level of a joint semantic system. However, for the specific type of gesture examined here, questions remain about whether the link between gesture and language reflects the engagement of an embodied conceptual-semantic network, a nonembodied symbolic process, or an interaction between the gesture modality input and amodal semantic representations in a semantic hub (Patterson, Nestor & Rogers, 2007; Caramazza & Masson, 2008). Given the degree of uncertainty still present in the literature, further studies are necessary to more definitively the processes determine involved in the links between gesture and lexical processing. One possibility would be the analysis of the cognitive processes and neural substrates that form the interface between gesture and action-verb representations by using, for example neuromodulation techniques (e.g. Transcranial Magnetic Stimulation (TMS).

As far as we are aware, this is the first study to analyse the link between pantomime recognition and action-verb processing in healthy subjects (for a study on action observation on verb retrieval difficulties in aphasia see Marangolo et al, 2010). We have demonstrated that verb production is facilitated by mere exposure to gestures, but that not all gestures prime verb production equally. This investigation supports a view of communication as multimodal, with language interacting naturally with gesture. The results not only inform theories of gesture-language interaction, but also suggest that gestures might be advantageously used to facilitate lexical retrieval in therapeutic interventions for language disorders.

References

Areshenkoff, C. N., Bub, D. N., & Masson M. E. J. (2015). *Task dependent motor representations evoked by spatial words: Implications for embodied accounts of word meaning*. Manuscript in preparation.

Bak, T. H., Yancopoulou, D., Nestor, P. J., Xuereb, J. H., Spillantini, M. G., Pulvermüller, F., & Hodges, J. R. (2006). Clinical, imaging and pathological correlates of a hereditary deficit in verb and action processing. *Brain, 129*(2), 321-332. doi:10.1093/brain/awh701.

Barry, C., Hirsh, K. W., Johnston, R. A., & Williams, C. L. (2001). Age of Acquisition, Word Frequency, and the Locus of Repetition Priming of Picture Naming. *Journal of Memory and Language, 44*(3), 350-375. doi:http://dx.doi.org/10.1006/jmla.2000.2743.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617(29).

Beattie, G., & Shovelton, H. (1999). Do iconic hand gestures really contribute anything to the semantic information conveyed by speech? An experimental investigation. *Semiotica*, *123* (1-2), 1-30. doi: 10.1515/semi.1999.123.1-2.1.

Bedny, M., Caramazza, A., Grossman, E., Pascual-Leone, A., & Saxe, R. (2008). Concepts Are More than Percepts: The Case of Action Verbs. *Journal of Neuroscience, 28*(44), 11347-11353. doi:10.1523/jneurosci.3039-08.2008.

Bernardis, P., & Carameli, N. (2009). Meaning in Words, Gestures and Mental Images. In N. Taatgen, & H. van Rijn, *Proceedings of the 31th Annual Conference of the Cognitive Science Society* (pp. 1692-1697). Cognitive Science Society.

Bernardis, P., Salillas, E., & Caramelli, N. (2008). Behavioural and neurophysiological evidence of semantic interaction between iconic gestures and words. *Cogn Neuropsychol, 25*(7-8), 1114-1128. doi:10.1080/02643290801921707.

Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V., & Rizzolatti, G. (2005). Listening to action-related sentences modulates the activity of the motor system: A combined TMS and behavioral study. *Cognitive Brain Research, 24*(3), 355-363.doi:10.1016/j.cogbrainres.2005.02.020.

Buxbaum, L. J., Kyle, K. M., & Menon, R. (2005). On beyond mirror neurons: Internal representations subserving imitation and recognition of skilled object-related actions in humans. *Cognitive Brain Research, 25*(1), 226-239. doi:http://dx.doi.org/10.1016/j.cogbrainres.2005.05.014.

Caramazza, A., Anzellotti, S., Strnad, L., Lingnau, A. (2014). Embodied Cognition and Mirror Neurons: A Critical Assessment. *Annual Review of Neuroscience, 37*(1), 1-15. doi:doi:10.1146/annurev-neuro-071013-013950.

Coelho, Y., & Bartolo, A. (2012). *Language and Action in Cognitive Neuroscience.* New York: NY: Psychology Press.

Cohen-Maximov, T., Avirame, K., Flöel, A., & Lavidor, M. (2015). Modulation of Gestural-verbal Semantic Integration by tDCS. *Brain Stimulation, 8*(3), 493-498. doi:10.1016/j.brs.2014.12.001.

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

Corballis, M. C. (2003). From mouth to hand: Gesture, speech, and the evolution of right-handedness. *Behavioral and Brain Sciences, 26*(02), 199-208. doi:10.1017/S0140525X03000062.

Cotelli, M., Borroni, B., Manenti, R., Alberici, A., Calabria, M., Agosti, C., . . . Cappa, S. F. (2006). Action and Object Naming in Frontotemporal Dementia, Progressive Supranuclear Palsy, and Corticobasal Degeneration. *Neuropsychology, 20*(5), 558-565. doi:10.1037/0894-4105.20.5.558.

Cotelli, M., Borroni, B., Manenti, R., Zanetti, M., Arévalo, A., Cappa, S. F., & Padovani, A. (2007). Action and object naming in Parkinson's disease without dementia. *European Journal of Neurology, 14*(6), 632-637.

De Aguiar, Vania (2015). *Enhancement of verb retrieval: Neuromodulation, repetition priming, and aphasia rehabilitation*. PhD thesis, University of Trento, University of Groningen, Macquarie University.

De Bleser, R., & Kauschke, C. (2003). Acquisition and loss of nouns and verbs: parallel or divergent patterns? *Journal of Neurolinguistics, 16*(2–3), 213-229. doi:http://dx.doi.org/10.1016/S0911-6044(02)00015-5.

De Marco, D., De Stefani, E., & Gentilucci, M. (2015). Gesture and word analysis: the same or different processes? *NeuroImage, 117*, 375-385. doi:10.1016/j.neuroimage.2015.05.080.

De Ruiter, J. (2000). The production of gesture and speech. In D. McNeill (Ed.), *Language and Gesture* (pp. 284-311). Cambridge: Cambridge University Press.

De Zubicaray, G., Arciuli, J., & McMahon, K. (2013). Putting an "end" to the motor cortex representations of action words. *Journal of Cognitive Neuroscencei, 25*(11), 1957-1974. doi:10.1162/jocn_a_00437.

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical Access in Aphasic and Nonaphasic Speakers. *Psychological Review, 104*(4), 801-838. doi:10.1037/0033-295X.104.4.801.

Dudschig, C., de la Vega, I., De Filippis, M., & Kaup, B. (2014). Language and vertical space: On the automaticity of language action interconnections. *Cortex, 58*, 151-160. doi:http://dx.doi.org/10.1016/j.cortex.2014.06.003.

Druks, J., & Masterson, J. (2000). *An Object and Action Naming Battery*. Hove: Psychology Press.

Ferro, J. M., Santos, M. E., Castro-caldas, A., & Mariano, M. G. (1980). Gesture recognition in aphasia. *Journal of Clinical Neuropsychology, 2*(4), 277-292. doi:10.1080/01688638008403800.

Feyereisen, P., Van de Wiele, M., & Dubois, F. (1988). The meaning of gestures: What can be understood without speech? *Cahiers de Psychologie Cognitive, 8*, 3-25.

Field, A. (2005). *Discovering Statistics using SPSS* (2nd ed.). London: Sage Publications.

Goldin-Meadow, S. (2003). *Hearing gesture: how our hands help us think.* Cambridge, MA: Cambridge, MA : Belknap Press of Harvard University Press.

Goldin-Meadow, S. (2007). Gesture with Speech and Without It. In S. Duncan, J. Cassel, & E. T. Levy (Eds.), *Gesture and the Dynamic Dimension of Language: Essays in honor of David McNeill* (pp. 99-108). Jonh Benjamins Publishing Company.

Goldin-Meadow, S., & Wagner, S. M. (2005). How our hands help us learn. *Trends in Cognitive Sciences, 9*(5), 234-241. doi:http://dx.doi.org/10.1016/j.tics.2005.03.006.

Hauk, O., Shtyrov, Y., & Pulvermuller, F. (2008). The time course of action and action-word comprehension in the human brain as revealed by neurophysiology. *Journal of Physiology, Paris, 102*(1-3), 50-58. doi:10.1016/j.jphysparis.2008.03.013.

Herrera, E., & Cuetos, F. (2012). Action naming in Parkinson's disease patients on/off dopamine. *Neuroscience Letters, 513*(2), 219-222. doi:http://dx.doi.org/10.1016/j.neulet.2012.02.045.

Herrera, E., Rodríguez-Ferreiro, J., & Cuetos, F. (2012). The effect of motion content in action naming by Parkinson’s disease patients. *Cortex, 48*(7), 900-904. doi:http://dx.doi.org/10.1016/j.cortex.2010.12.007.

Hostetter, A., & Alibali, M. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin & Review, 15*(3), 495-514. doi:10.3758/PBR.15.3.495.

Hostetter, A. B., & Alibali, M. W. (2010). Language, Gesture, Action! A Test of the Gesture as Simulated Action Framework. *Journal of Memory and Language, 63*(2), 245-257. doi:10.1016/j.jml.2010.04.003.

Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: experimental and computational studies. *Cognition, 100*(3), 464-482. doi:http://dx.doi.org/10.1016/j.cognition.2005.02.006.

Innocenti, A., De Stefani, E., Sestito, M., & Gentilucci, M. (2014). Understanding of action-related and abstract verbs in comparison: a behavioral and TMS study. *Cognitive Processing, 15*(1), 85-92. doi:10.1007/s10339-013-0583-z..

Jackendoff, R. (1990). *Semantic Structures*. MIT Press, Cambridge, MA.

Jonkers, R. (1998). *Comprehension and production of verbs in aphasic speakers*. Groningen: s.n.

Jonides, J., & Mack, R. (1984). On the cost and benefit of cost and benefit. *Psychological Bulletin, 96*(1), 29-44. doi:10.1037/0033-2909.96.1.29.

Jonkers, R., & Bastiaanse, R. (2007). Action naming in anomic aphasic speakers: Effects of instrumentality and name relation. *Brain and Language, 102*(3), 262-272. doi:http://dx.doi.org/10.1016/j.bandl.2007.01.002.

Kauschke, C., & von Frankenberg, J. (2008). The differential influence of lexical parameters on naming latencies in German. A study on noun and verb picture naming. *Journal Of Psycholinguistic Research, 37*(4), 243-257.

Kelly, S. D., Özyürek , A. & Maris, E. (2010). Two Sides of the Same Coin: Speech and Gesture Mutually Interact to Enhance Comprehension. *Psychological Science, 21*(2), 260-267.

Kemmerer, D. (2015). Visual and motor features of the meanings of action verbs: A cognitive neuroscience perspective. In R. de Almeida, & C. Manouilidou (Eds.), *Cognitive Science Perspectives on Verb Representation and Processing* (pp. 189-205). Switzerland: Springer International Publishing.

Kemmerer, D., & Gonzalez-Castillo, J. (2010). The Two-Level Theory of verb meaning: An approach to integrating the semantics of action with the mirror neuron system. *Brain and Language, 112*(1), 54-76. doi:10.1016/j.bandl.2008.09.010.

Kemmerer, D., Rudrauf, D., Manzel, K., & Tranel, D. (2012). Behavioral patterns and lesion sites associated with impaired processing of lexical and conceptual knowledge of actions. *Cortex, 48*(7), 826-848. doi:10.1016/j.cortex.2010.11.001.

Kipper-Schuler, K. (2006). *VerbNet: A broad-coverage, comprehensive verb lexicon*. PhD thesis. University of Pennsylvania, Filadelfia.

Kiss, G.R., Armstrong, C., Milroy, R., Piper, J. (1973) An associative thesaurus of English and its computer analysis. In Aitken AJ, Bailey RW, Hamilton-Smith N. (Eds). *The computer and literary studies.* Edinburgh: Edinburgh University Press. pp. 153–165.

Kita, S. (2000). How representational gestures help speaking. In D. McNeill (Ed.), *Language and Gesture* (pp. 162-185). Cambridge: Cambridge University Press.

Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language, 48*(1), 16-32. doi:10.1016/S0749-596X(02)00505-3.

Krauss, R., Chen, Y., & Gottesman, R. (2000). Lexical gestures and lexical access: A process model. In D. McNeill (Ed.), *Language and gesture* (pp. 261-283). New York: Cambridge University Press.

Levelt, W. J. M. (1989). *Speaking: from intention to articulation.* Cambridge, Mass: MIT Press.

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

Luzzatti, C., Raggi, R., Zonca, G., Pistarini, C., Contardi, A., Pinna, G. D. (2002). Verb-noun double dissociation in aphasic lexical impairments: the role of word frequency and imageability. *Brain and Language*, 444(81), 432–444. doi:10.1006/brln.2001.2536.

Mahon, B. Z., & Caramazza, A. (2009). Concepts and Categories: A Cognitive Neuropsychological Perspective. *Annual Review of Psychology, 60*(1), 27-51. doi:10.1146/annurev.psych.60.110707.163532.

Maieron, M., Marin, D., Fabbro, F., & Skrap, M. (2013). Seeking a bridge between language and motor cortices: a PPI study. *Frontiers in Human Neurosciences, 7*, 249. doi:10.3389/fnhum.2013.00249.

Marangolo, P., Bonifazi, S., Tomaiuolo, F., Craighero, L., Coccia, M., Altoe, G., . . . Cantagallo, A. (2010). Improving language without words: first evidence from aphasia. *Neuropsychologia, 48*(13), 3824-3833. doi:10.1016/j.neuropsychologia.2010.09.025.

Masson, M. J., Bub, D. N., & Newton-Taylor, M. (2008). Language-based access to gestural components of conceptual knowledge. *The Quarterly Journal of Experimental Psychology, 61*(6), 869-882. doi:10.1080/17470210701623829.

Masson-Carro, I., Goudbeek, M., & Krahmer, E. (2015). Coming of age in gesture: A comparative study of gesturing and pantomiming in older children and adults. In *Proceedings of the 4th GESPIN - Gesture & Speech in Interaction Conference.*

McNamara, T. P. (2005). *Semantic priming : perspectives from memory and word recognition.* New York: Psychology Press.

McNeill, D. (1992). *Hand and mind : what gestures reveal about thought / David McNeill*. Chicago: Chicago : University of Chicago Press.

McNeill , D. (2000). *Language and Gesture.* Cambridge: Cambridge University Press.

McNeill, D. (2006). Gestures and communication. In K. Brown (*2nd Ed.*), *Encyclopedia of Language and Linguistics* (pp.58-66).

McNeill, D., Cassell, J., & McCullough, K. E. (1994). Communicative Effects of Speech-Mismatched Gestures. *Research on Language and Social Interaction, 27*(3), 223-237. doi:10.1207/s15327973rlsi2703_4.

McRae, K., De Sa, V. R., & Seidenberg, M. S. (1997). On the Nature and Scope of Featural Representations of Word Meaning. *Journal of Experimental Psychology: General, 126*(2), 99-130. doi:10.1037/0096-3445.126.2.99.

Morrel-Samuels, P., & Krauss, R. M. (1992). Word Familiarity Predicts Temporal Asynchrony of Hand Gestures and Speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(3), 615-622. doi:10.1037/0278-7393.18.3.615.

Moseley, R. L., & Pulvermüller, F. (2014). Nouns, verbs, objects, actions, and abstractions: Local fMRI activity indexes semantics, not lexical categories. *Brain and Language, 132*(100), 28-42. doi:10.1016/j.bandl.2014.03.001.

Moy, G., & Mounoud, P. (2003). Object recognition in young adults: is priming with pantomimes possible? In *Catalogue des Abstracts: 8ème Congrès de la Société Suisse de Psychologie (SSP)*: Berne.

Mounoud, P., Duscherer, K., Moy, G., & Perraudin, S. (2007). The influence of action perception on object recognition: a developmental study. *Developmental Science, 10*(6), 836-852. doi:10.1111/j.1467-7687.2007.00624.x.

Muslimovic, D., Post, B., Speelman, J. D., & Schmand, B. (2005). Cognitive profile of patients with newly diagnosed Parkinson disease. *Neurology, 65*(8), 1239-45.

Neely, J. (1991). Semantic priming: effects in visual word recognition: a selective review of current findings and theories. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading: visual and word recognition*. Hillsdale, N.J.: L. Erlbaum Associates.

Negri, G. A., Rumiati, R. I., Zadini, A., Ukmar, M., Mahon, B. Z., & Caramazza, A. (2007). What is the role of motor simulation in action and object recognition? Evidence from apraxia. *Cognitive Neuropsychology, 24*(8), 795-816. doi:10.1080/02643290701707412.

Nickels, L. (2001). Spoken word production. In B. Rapp (Ed.), *The handbook of cognitive neuropsychology* (pp. 291-320). New York: Psychology Press.

Nispen, K. v., Sandt-Koenderman, M. v., Mol, L., & Krahmer, E. (2014). Pantomime Strategies: On Regularities in How People Translate Mental Representations into the Gesture Modality. In P. Bello,

M. Guarini, M. McShane, & B. Scassellati, *Proceedings of the 36th Annual Conference of the Cognitive Science Society (CogSci 2014)* (pp. 3020-3025). Austin, TX: Cognitive Science Society.

O'Reilly, A. W. (1995). Using Representations: Comprehension and Production of Actions with Imagined Objects. *Child Development, 66*(4), 999-1010. doi:10.1111/j.1467-8624.1995.tb00918.x.

Oliveri, M., Finocchiaro, C., Shapiro, K., Gangitano, M., Caramazza, A., & Pascual-Leone, A. (2004). All Talk and No Action: A Transcranial Magnetic Stimulation Study of Motor Cortex Activation during Action Word Production. *Journal of Cognitive Neuroscience, 16*(3), 374-381. doi:10.1162/089892904322926719.

Osiurak, F., Jarry, C., Baltenneck, N., Boudin, B., & Le Gall, D. (2012). Make a gesture and I will tell you what you are miming. Pantomime recognition in healthy subjects. *Cortex, 48*(5), 584-592. doi:10.1016/j.cortex.2011.01.007.

Özyürek, A. (2014). Hearing and seeing meaning in speech and gesture: insights from brain and behaviour. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 369*(1651), 20130296. doi:10.1098/rstb.2013.0296.

Papagno, C., Mattavelli, G., Cattaneo, Z., Romito, L., & Albanese, A. (2013). Ambiguous idiom processing in Parkinson's disease patients. *Cognitive Neuropsychology, 30*(7-8), 495-506. doi:10.1080/02643294.2013.876397.

Papeo, L., Vallesi, A., Isaja, A., & Rumiati, R. I. (2009). Effects of TMS on Different Stages of Motor and Non-Motor Verb Processing in the Primary Motor Cortex. *PLoS ONE, 4*(2), e4508. doi:10.1371/journal.pone.0004508.

Patterson, K., Nestor, P., J. , & Rogers, T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience, 8*(12), 976. doi:10.1038/nrn2277.

Perniss, P., Thompson, R., & Vigliocco, G. (2010). Iconicity as a general property of language: evidence from spoken and signed languages *Frontiers in Psychology, 1,* 227. doi:10.3389/fpsyg.2010.00227.

Perniss, P., & Vigliocco, G. (2014). The bridge of iconicity: from a world of experience to the experience of language. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 369*(1651), 20130300. doi:10.1098/rstb.2013.0300.

Pine, K. J., Reeves, L., Howlett, N., & Fletcher, B. C. (2013). Giving cognition a helping hand: the effect of congruent gestures on object name retrieval. *British Journal of Psychology, 104*(1), 57-68. doi:10.1111/j.2044-8295.2011.02098.x.

Pinker, S. (1989). *Learnability and cognition: the acquisition of argument structure*. Cambridge, MA: MIT Press.

Plant, C., Webster, J., & Whitworth, A. (2011). Category norm data and relationships with lexical frequency and typicality within verb semantic categories. *Behavior Research Methods, 43*(2), 424-440. doi:10.3758/s13428-010-0051-y.

Plaut, D. C. (1995). Semantic and associative priming in a distributed attractor network. *Proceedings of the 17th Annual Conference of the Cognitive Science Society* (pp. 37-42). Hillsdale, NJ: Lawrence Erlbaum Associates.

Pulvermüller, F., Hauk, O., Nikulin, V. V., & Ilmoniemi, R. J. (2005). Functional links between motor and language systems. *European Journal of Neuroscience, 21*(3), 793-797. doi:10.1111/j.1460-9568.2005.03900.x.

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. doi:http://dx.doi.org/10.1016/j.neuropsychologia.2012.11.001.

Rappaport, M., & Levin, B. (1988). What to do with theta-roles. In W. Wilkins (Ed.) *Syntax and Semantics, Vol. 21: Thematic Relations*, Academic Press, San Diego, CA.

Rose, M. L. (2006). The utility of arm and hand gestures in the treatment of aphasia. *International Journal of Speech-Language Pathology, 8*(2), 92-109. doi:10.1080/14417040600657948.

Rothi, L. J., Heilman, K. M., & Watson, R. T. (1985). Pantomime comprehension and ideomotor apraxia. *Journal of Neurology, Neurosurgery & Psychiatry, 48*(3), 207. doi:10.1136/jnnp.48.3.207.

Rothi, L.J., Ochipa, C. & Heilman, K.M. (1997). A Cognitive Neuropsychological Model of Limb Praxis and Apraxia. In L.J. Rothi and K.M. Heilman (Eds.) *Apraxia: the neuropsychology of action* (pp. 29-49). Hove: Hove : Psychology.

Rumiati, R. I., & Humphreys, G. W. (1998). Recognition by Action: Dissociating Visual and Semantic Routes to Action in Normal Observers. *Journal of Experimental Psychology: Human Perception and Performance, 24*(2), 631-647. doi:10.1037/0096-1523.24.2.631.

Saygin, A. P., Wilson, S. M., Dronkers, N. F., & Bates, E. (2004). Action comprehension in aphasia: linguistic and non-linguistic deficits and their lesion correlates. *Neuropsychologia, 42*(13), 1788-1804. doi:10.1016/j.neuropsychologia.2004.04.016.

Scharp, V. L., Tompkins, C. A., & Iverson, J. M. (2007). Gesture and aphasia: Helping hands? *Aphasiology, 21*(6-8), 717-725. doi:10.1080/02687030701192273.

Silveri, M. C., Salvigni, B. L., Cappa, A., Della Vedova, C., & Puopolo, M. (2003). Impairment of Verb Processing in Frontal Variant-Frontotemporal Dementia: A Dysexecutive Symptom. *Dementia and Geriatric Cognitive Disorders, 16*(4), 296-300.

Ska, B., & Croisile, B. (1998). Effects of normal aging on the recognition of gestures. *Brain and Cognition., 37*(1), 136-138.

So, W. C., Yi-Feng, A. L., Yap, D. F., Kheng, E., & Yap, J. M. (2013). Iconic gestures prime words: comparison of priming effects when gestures are presented alone and when they are accompanying speech. *Frontiers in Psychology, 4*, 779. doi:10.3389/fpsyg.2013.00779.

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., . . . Bates, E. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language, 51*(2), 247-250. doi:http://dx.doi.org/10.1016/j.jml.2004.03.002.

Tabachnick, B. G. & Fidell, L. (2013). *Using multivariate statistics* (6th edition): Boston Pearson.

Thompson, C. K., Lange, K. L., Schneider, S. L., & Shapiro, L. P. (1997). Agrammatic and non-brain-damaged subjects' verb and verb argument structure production. *Aphasiology, 11*(4-5), 473-490. doi:10.1080/02687039708248485.

Tomasino, B., Ceschia, M., Fabbro, F., & Skrap, M. (2012). Motor Simulation during Action Word Processing in Neurosurgical Patients. *Journal of Cognitive Neuroscience, 24*(3), 736-748. doi:10.1162/jocn_a_00168.

Tomasino, B., Ceschia, M., Fabbro, F., & Skrap, M. (2012). Motor Simulation during Action Word Processing in Neurosurgical Patients. *Journal of Cognitive Neuroscience*, 24(3), 736-748. doi:10.1162/jocn_a_00168.

Tomasino, B., & Rumiati, R. I. (2013). At the mercy of strategies: the role of motor representations in language understanding. *Frontiers in Psychology, 4*, 27. doi:10.3389/fpsyg.2013.00027.

Vigliocco, G., Vinson, D. P., Lewis, W., & Garrett, M. F. (2004). Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive Psychology, 48*(4), 422-488. doi:http://dx.doi.org/10.1016/j.cogpsych.2003.09.001.

Vinson, D., Thompson, R. L., Skinner, R., & Vigliocco, G. (2015). A faster path between meaning and form? Iconicity facilitates sign recognition and production in British Sign Language. *Journal of Memory and Language, 82*, 56-85. doi:10.1016/j.jml.2015.03.002.

Vinson, D. P., & Vigliocco, G. (2002). A semantic analysis of grammatical class impairments: semantic representations of object nouns, action nouns, and action verbs. *Journal of Neurolinguistics, 15*, 317–351.

Vinson, D. P., Vigliocco, G., Cappa, S., & Siri, S. (2003). The breakdown of semantic knowledge: Insights from a statistical model of meaning representation. *Brain and Language, 86*(3), 347-365. doi:http://dx.doi.org/10.1016/S0093-934X(03)00144-5.

Vitkovitch, M., Rutter, C., & Read, A. (2001). Inhibitory effects during object name retrieval: the effect of interval between prime and target on picture naming responses. *British Journal of Psychology, 92*(Pt 3), 483-506.

Wang, L., & Goodglass, H. (1992). Pantomime, praxis, and aphasia. *Brain and Language, 42*(4), 402-418. doi:10.1016/0093-934X(92)90076-Q.

Watson, C. E., Cardillo, E. R., Ianni, G. R., & Chatterjee, A. (2013). Action Concepts in the Brain: An Activation Likelihood Estimation Meta-analysis. *Journal of Cognitive Neuroscience, 25*(8), 1191-1205. doi:10.1162/jocn_a_00401.

Wheeler, K. B., & Bergen, B. K. (2006) Meaning in the palm of your hand. In S. Rice & J. Newman (Eds.), *Empirical and experimental methods in conceptual structure, discourse and language* (pp. 2–13). Stanford: CSLI.

Wheeldon, L., & Monsell, S. (1992) The Locus of Repetition Priming of Spoken Word Production. *The Quarterly Journal of Experimental Psychology*, 44 (4), 723-761, doi: 10.1080/14640749208401307.

Wheeldon, L. R., & Monsell, S. (1994). Inhibition of Spoken Word Production by Priming a Semantic Competitor. *Journal of Memory and Language, 33*(3), 332-356. doi:10.1006/jmla.1994.1016.

Willems, R. M., & Casasanto, D. (2011). Flexibility in Embodied Language Understanding. *Frontiers in Psychology, 2*. doi:10.3389/fpsyg.2011.00116.

Willems, R. M., Özyürek, A., & Hagoort, P. (2007). When language meets action: the neural integration of gesture and speech. *Cereb Cortex, 17*(10), 2322-2333. doi:10.1093/cercor/bhl141.

Yang, J., Andric, M., & Mathew, M. M. (2015). The neural basis of hand gesture comprehension: A meta-analysis of functional magnetic resonance imaging studies. *Neuroscience & Biobehavioral Reviews, 57*, 88-104. doi:http://dx.doi.org/10.1016/j.neubiorev.2015.08.006.

Yap, D. F., So, W. C., Yap, J. M., Tan, Y. Q., & Teoh, R. L. (2011). Iconic gestures prime words. *Cognitive Science, 35*(1), 171-183. doi:10.1111/j.1551-6709.2010.01141.x.

APPENDIX A: Gesture name-appropriateness rating, correct responses and proportion of error types for each gesture

Table A 1
Gesture name-appropriateness rating, correct responses and proportion of error types for each gesture

item	mean name-appropriateness rating	Naming							
		Prop. of Correct responses	Prop. of error types (total responses)			Prop. of errors	Prop. of error types (total errors)		
			Semantic errors	Visual errors	Other errors		Semantic errors	Visual errors	Other errors
salute	4.93	1	0	0	0	0	0	0	0
sleep	4.86	1	0	0	0	0	0	0	0
drink	4.86	1	0	0	0	0	0	0	0
whistle	4.79	0.90	0	0.10	0	0.10	0	1	0
knock	4.79	1	0	0	0	0	0	0	0
hug	4.76	1	0	0	0	0	0	0	0
clap	4.76	0.93	0.07	0	0	0.07	1	0	0
smoke	4.66	0.93	0	0.07	0	0.07	0	1	0
wave	4.62	0.97	0	0	0.03	0.03	0	0	1
scratch	4.62	0.97	0.03	0	0	0.03	1	0	0
pray	4.62	1	0	0	0	0	0	0	0
listen	4.62	0.97	0.03	0	0	0.03	1	0	0
cry	4.62	1	0	0	0	0	0	0	0
push	4.59	0.83	0	0.07	0.1	0.17	0	0.41	0.59
throw	4.59	0.87	0.13	0	0	0.13	1	0	0
drive	4.59	0.80	0.2	0	0	0.20	1	0	0
type	4.55	0.93	0	0	0.07	0.07	0	0	1
sneeze	4.55	0.97	0.03	0	0	0.03	1	0	0
zip	4.52	0.90	0	0.10	0	0.10	0	1	0

item	mean name-appropriateness rating	Naming							
		Prop. of Correct responses	Prop. of error types (total responses)			Prop. of errors	Prop. of error types (total errors)		
			Semantic errors	Visual errors	Other errors		Semantic errors	Visual errors	Other errors
comb	4.52	0.47	0.53	0	0	0.53	1	0	0
chop	4.52	0.80	0.20	0	0	0.20	1	0	0
shoot	4.45	0.77	0.03	0.17	0.03	0.23	0.13	0.74	0.13
dive	4.45	0.83	0.03	0.13	0	0.16	0.19	0.81	0
conduct	4.31	0.53	0.23	0.23	0	0.46	0.50	0.50	0
shave	4.31	0.90	0	0.03	0.07	0.10	0	0.30	0.70
pour	4.28	0.93	0	0.07	0	0.07	0	1	0
walk	4.28	0.93	0.03	0.03	0	0.06	0.50	0.50	0
twist	4.24	0.23	0.73	0.03	0	0.76	0.96	0.04	0
spread	4.21	0.27	0.23	0.27	0.23	0.73	0.31	0.37	0.32
sweat	4.17	0.50	0	0.50	0	0.50	0	1	0
shiver	4.17	0.47	0.07	0.17	0.30	0.54	0.13	0.31	0.56
wash	4.14	0.87	0.13	0	0	0.13	1	0	0
stir	4.14	0.67	0.3	0.03	0	0.33	0.91	0.09	0
break	4.1	0.67	0.17	0.10	0.07	0.34	0.50	0.29	0.21
cut	4.1	0.87	0.10	0	0.03	0.13	0.77	0	0.23
yell	4.07	0.53	0.33	0.03	0.1	0.46	0.72	0.07	0.22
eat	4.03	0.87	0	0.13	0	0.13	0	1	0
fish	4	0.40	0.47	0.10	0.03	0.60	0.78	0.17	0.05
blow	4	0.77	0.03	0.10	0.10	0.23	0.13	0.44	0.43
open	4	0.90	0.03	0.07	0	0.10	0.30	0.70	0
swim	4	1	0	0	0	0	0	0	0

item	mean name-appropriateness rating	Naming							
		Prop. of Correct responses	Prop. of error types (total responses)			Prop. of errors	Prop. of error types (total errors)		
			Semantic errors	Visual errors	Other errors		Semantic errors	Visual errors	Other errors
tie	3.97	0.87	0	0.03	0.10	0.13	0	0.23	0.77
swing	3.86	0.70	0.13	0.07	0.10	0.30	0.43	0.23	0.33
brush	3.86	0.90	0.07	0.03	0	0.10	0.70	0.3	0
climb	3.86	1	0	0	0	0	0	0	0
slide	3.83	0.10	0	0.90	0	0.90	0	1	0
sew	3.79	0.77	0.10	0.07	0.07	0.24	0.42	0.29	0.29
write	3.76	0.37	0	0.53	0.10	0.63	0	0.84	0.16
squeeze	3.76	0.23	0	0.47	0.30	0.77	0	0.61	0.39
hammer	3.76	0.53	0.07	0.40	0	0.47	0.15	0.85	0
ski	3.69	0.47	0	0.43	0.10	0.53	0	0.81	0.19
give	3.69	0.67	0	0.30	0.03	0.33	0	0.91	0.09
peel	3.69	0.57	0.03	0.17	0.23	0.43	0.07	0.40	0.53
bowl	3.69	0.57	0.40	0	0.03	0.43	0.93	0	0.07
saw	3.69	0.37	0.60	0	0.03	0.63	0.95	0	0.05
read	3.62	0.23	0.30	0.17	0.30	0.77	0.39	0.22	0.39
bounce	3.55	0.30	0.23	0.40	0.07	0.70	0.33	0.57	0.1
watch	3.55	0.13	0.73	0.07	0.07	0.87	0.84	0.08	0.08
iron	3.48	0.17	0.10	0.47	0.27	0.84	0.12	0.56	0.32
erase	3.38	0.33	0	0.60	0.07	0.67	0	0.90	0.10
tear	3.34	0.37	0.50	0.13	0	0.63	0.79	0.21	0
row	3.31	0.40	0.07	0.50	0.03	0.60	0.12	0.83	0.05
paint	3.17	0.60	0.03	0.27	0.10	0.40	0.08	0.67	0.25
pull	3.17	0.83	0.07	0.07	0.03	0.17	0.42	0.41	0.18

item	mean name-appropriateness rating	Naming							
		Prop. of Correct responses	Prop. of error types (total responses)			Prop. of errors	Prop. of error types (total errors)		
			Semantic errors	Visual errors	Other errors		Semantic errors	Visual errors	Other errors
stroke	3	0.03	0.37	0.47	0.13	0.97	0.38	0.49	0.13
light	3	0.17	0.33	0.27	0.23	0.83	0.40	0.32	0.28
run	3	0.70	0.13	0.03	0.13	0.29	0.45	0.10	0.45
clean	2.97	0.07	0.33	0.47	0.13	0.93	0.35	0.51	0.14
rake	2.93	0.10	0.13	0.40	0.37	0.90	0.14	0.45	0.41
box	2.9	0.20	0.80	0	0	0.80	1	0	0
vacuum	2.83	0	0.37	0.63	0	1	0.37	0.63	0
play	2.83	0.43	0	0.50	0.07	0.57	0	0.88	0.12
whisper	2.76	0.57	0.37	0.03	0.03	0.43	0.86	0.07	0.07
cook	2.72	0	0.37	0.57	0.07	1.01	0.37	0.56	0.07
pay	2.66	0.20	0.17	0.13	0.50	0.80	0.21	0.16	0.63
knit	2.52	0.10	0	0.33	0.57	0.90	0	0.37	0.63
carry	2.38	0.07	0	0.67	0.27	0.94	0	0.71	0.29
steal	2.38	0	1	0	0	1	1	0	0
talk	2.25	0.23	0.17	0.33	0.27	0.77	0.22	0.43	0.35
skip	2.21	0.37	0.70	0.07	0.23	1	0.70	0.07	0.23
dance	2.17	0.27	0.07	0.67	0	0.74	0.09	0.91	0
mop	2.14	0.03	0	0.33	0.63	0.96	0	0.34	0.66
hang	2.07	0.03	0.47	0.20	0.30	0.97	0.48	0.21	0.31
sweep	2	0.07	0.13	0.40	0.40	0.93	0.14	0.43	0.43
water	1.79	0.03	0.07	0.53	0.37	0.97	0.07	0.55	0.38
dust	1.62	0.17	0.23	0.33	0.27	0.83	0.28	0.40	0.32

APPENDIX B: Priming effects for each of the target verbs

Table B 1

Items showing faster naming latency in the match condition (ordered by size of priming effect)

Item	Match (mean RT ms)	*std*	Mismatch (mean RT ms)	*std*	Priming effect (ms)	**t-test (two-tailed)**
clap	623	92	1134	353	511	*p* <.001**
wave	756	217	1265	285	510	*p* <.001**
knock	673	117	1082	283	408	*p* <.001**
scratch	745	168	1096	195	351	*p* <.001**
whistle	770	210	1118	302	348	*p* <.001**
wash	919	141	1146	241	328	*p* <.001**
break	905	169	1223	235	318	*p* <.001**
cry	699	130	1010	333	312	*p* <.001**
sneeze	620	133	921	304	301	*p* <.001**
spread	950	245	1284	295	298	*p* <.001**
throw	762	196	1044	247	282	*p*= .016*
give	961	144	1231	229	270	*p* <.001**
zip	794	286	1056	345	263	*p* <.001**
light	1084	268	1346	333	262	*p* <.001**
chop	950	161	1226	306	276	*p*= .001**
walk	767	846	1022	217	256	*p* <.001**
erase	1067	204	1320	247	252	*p* <.001**
pull	957	221	1201	310	244	*p* <.001**
drive	689	126	930	197	242	*p* <.001**
open	829	187	1069	246	239	*p* <.001**
shoot	674	184	908	213	233	*p* <.001**
watch	952	208	1185	289	233	*p* <.001**
tear	757	149	980	224	223	*p*= .020*
swim	592	148	794	153	201	*p* <.001**
dive	746	155	943	225	197	*p* <.001**
run	768	211	960	142	193	*p* <.001**
climb	940	215	1131	332	191	*p* <.001**
tie	831	133	1015	204	184	*p* <.001**
pour	821	257	1002	216	181	*p* <.001**
smoke	613	120	792	142	179	*p* <.001**
bowl	854	232	1012	188	158	*p* <.001**
hug	729	174	882	223	153	*p* <.001**
shave	755	198	904	139	149	*p* <.001**
shiver	833	258	976	147	143	*p*= .002**

Item	Match (mean RT ms)	*std*	Mismatch (mean RT ms)	*std*	Priming effect (ms)	**t-test (two-tailed)**
salute	809	242	941	270	132	*p* <.001**
conduct	962	244	1092	200	130	*p* <.001**
type	763	160	887	170	124	*p* <.001**
whisper	945	145	1067	190	122	*p* <.001**
row	783	215	899	180	116	*p* <.001**
swing	749	171	865	203	116	*p*= .197
skip	731	236	842	272	110	*p* <.001**
clean	1211	203	1319	236	107	*p* <.001**
paint	853	204	959	170	106	*p* <.001**
drink	721	206	808	169	87	*p*= .001**
yell	933	233	1012	205	77	*p*= .003**
fish	918	208	982	169	64	*p*= .001**
box	956	141	1019	209	63	*p*= .027*
sleep	683	249	745	97	62	*p*= .136
peel	940	103	995	270	54	*p*= .319
write	905	248	958	157	54	*p*= .083
rake	942	149	991	224	49	*p*= .195
push	840	172	873	186	33	*p*= .005**
bounce	1009	263	1033	264	24	*p*= .861
vacuum	927	221	948	251	21	*p*= .304
cut	914	213	934	256	20	*p*= .451
eat	916	224	932	158	16	*p*= .864
dance	984	157	993	159	9	*p*= .552
cook	1075	215	1078	212	3	*p*= .873
pray	835	153	820	114	-15	*p*= .272
ski	807	190	786	136	-22	*p*= .235
brush	916	248	890	183	-26	*p*= .208
play	1191	224	1154	355	-37	*p*= .393
blow	934	184	891	148	-43	*p*= .003**
sew	1120	222	1075	198	-45	*p*= .037*
hammer	1001	383	955	279	-45	*p*= .301
squeeze	1016	322	970	242	-46	*p*= .169
saw	1002	303	924	122	-78	*p*= .243
sweat	985	272	834	74	-91	*p*= .164
slide	885	207	770	103	-115	*p*= .002**

* difference between match and mismatch conditions is significant at 0.05 level; ** difference between match and mismatch conditions is significant at 0.01 level