1 'I Am Trying to Climb Everest in Flip-Flops.' What Is Embodied Metaphor and Where Does It Come From?

1.1 Introduction

Metaphor is a device through which we perceive or experience one entity in terms of another, but what is 'embodied metaphor' and where does it come from? In order to begin to answer this question, I would like to consider the following extract from a blog written by a woman who had recently experienced the loss of her baby, at birth:

> This process of *moving through* his death, and the life we had already *built* around his too-short existence, is our own to deal with; I know this. But like all of us, *I look for insight*; *some key that will unlock the endless hallway of doors now closed in front of me.*
>
> The problem is, *I don't actually look at grief as a series of doors to unlock over time*. Rather, *it feels much more like I'm standing behind a wall.*
>
> *It's a wall without a door and without a key. There are no keys* in the online forums, the medical and mental health websites, from the grief counsellors, in the support groups, or from the wisdom of family or friends. No one tells you what to do *when you lose yourself.*
>
> *What I hold onto*, I guess … is the hope that I'll *wake up* and *the wall will have finally crumbled.* https://whatsyourgrief.com/loss-of-identity-after-stillbirth/

As we can see in the italicised parts of this extract, the writer makes substantial use of metaphor to describe her experience. She talks about '*moving through* his death', her need to '*unlock* the endless hallway of doors', the fact that she '*feels'* like she is '*standing* behind a wall', her hope that she will one day '*wake up'* from this nightmare. What is striking about all of these metaphors is the fact that they involve bodily actions and senses. It is as if she is not only describing her grief through these metaphors, but that she is actually, on one level, 'experiencing' it through these metaphors.

Metaphors that draw on physical experiences can be found in all languages. They are not uncommon, and many are highly conventional. For example, in English, 'understanding' is often talked about in terms of 'seeing' ('I can *see* why Jim is confused'[1]), 'time' is often talked about in terms of 'space' (e.g. 'He had *come a long way* since his early days as a security guard'), achieving

professional advancement is often talked about in terms of upward movement (e.g. 'He *rose* through the ranks to become head inspector'). Many metaphors make reference to the senses, allowing people to say, for example, that an experience 'leaves a *bad taste* in their mouths', that '*sharp* words were exchanged' or that one can '*smell* out bad practice'.

It is now widely accepted that metaphors such as these shape our understanding of the world. They are part of our conceptual system, and without them, we would find it virtually impossible to reason and communicate about abstract concepts. For example, we think about affection as warmth, importance as size, intimacy as closeness, difficulties as burdens, categories as containers, similarity as closeness, organisation as physical structure, time as motion, purposes as destinations, causes as physical forces, knowledge as sight and understanding as grasping (Lakoff and Johnson, 1999). In each of these cases, an abstract concept is understood in more concrete, physical terms via metaphor.

There is now a substantial body of research showing that metaphors such as these are not just part of our language but that at times they are actually 'experienced' albeit at a subconscious level. In other words, when we encounter them, subconscious sensorimotor responses are triggered that are similar to those that would be triggered if we actually observed or experienced these actions and senses in the physical world (Gibbs, 2006a). These metaphors therefore have the potential to be experienced on a physical level, rather than being purely an external, objective phenomenon. Gibbs (2006b) refines this argument by suggesting that part of our ability to make sense of metaphors such as these 'resides in the automatic construction of a simulation whereby we imagine performing the bodily actions referred to in these excerpts' (ibid., 435). He goes on to argue that key areas of the brain, such as the motor cortex, are involved in the processing of metaphors such as these but 'as importantly, people's intuitive, felt, phenomenological experiences of their own bodies shape large portions of metaphoric thought and language use' (ibid., 436). For this reason, metaphors such as these are often described as 'embodied', though the degree of embodiment can range from full-on sensorimotor activation through to the use of bodily knowledge in shaping our metaphorical thought processes. The fact that metaphors can be experienced in this way is not a trivial observation for, as we have just seen, physical human experiences form the basis of a very large number of metaphors, and most, if not all, abstract phenomena are understood through metaphors, many of which are based on bodily experiences.

The majority of humans inhabit roughly the same types of bodies and use them to do roughly the same sorts of things, so there is a great deal of universality in the ways in which we experience metaphor. For this reason, as we will see in Chapter 2, much of the literature on embodied metaphor has

focused on the homogeneous nature of human experience and has de-emphasised possible sources of variation. It has been argued that many of the aforementioned metaphorical relationships are universal because they are based on correlations between bodily experiences such as that between affection and warmth or visual perception and understanding, and these are experienced by all humankind (Grady, 1997a, b). There has been a substantial amount of work on cross-cultural variation (e.g. Kövecses, 2015; Musolff, 2017) but the impact of other sources of variation, such as gender, personality, body shape and disability has received very little attention. Notions of universality are prevalent in the world of embodied metaphor.

The aim of this book is to challenge these notions of universality by identifying and exploring sources of variation in human experience that affect the ways in which people make sense of the world through embodied metaphor. I take the position that because people's experiences with the physical world differ in important ways, in physical, emotional and social terms, our experiences of embodied metaphor are also susceptible to variation. By understanding these sources of variation, we will gain deeper insights into the different types of world views that people develop, why they develop them and the mechanisms through which they develop them. We can also use this knowledge to further our understanding of embodied metaphor, or, in other words, what it means to 'experience' the world through metaphor.

1.2 Embodied Cognition

The idea that metaphor is embodied rests on the more fundamental premise that much of cognition itself is embodied. According to the embodied cognition hypothesis (Rosch et al., 1991), the ways in which we interact through our bodies with the world and with people around us serve as a basis for the way in which we form ideas and communicate these ideas to others. In other words, our perceptual, motor and other experiences play a fundamental role in how we talk about, think about and interact with people, objects and the world in general. In Gibbs' words: 'people's subjective, felt experience of their bodies in action provides part of the fundamental grounding for language and thought' (Gibbs, 2006a: 2). Stimuli are thus not just seen or heard, but 'experienced' in the body, leading to the activation of parts of the brain that are involved in movement, processing of sensory stimuli and emotion. So, for example, when we watch a footballer score a deciding goal in the final minute of a football match and then express his jubilance by sliding through the mud on his knees, we mentally simulate both his actions and his state of mind as if they were actually happening (Gallese, 2006) and this is the reason why we empathise with and enjoy the experience.

The embodied cognition hypothesis is inspired, at least in part, by Barsalou's (1999) perceptual systems theory, according to which the simple perception of a stimulus leads to the simulation of actual perceptual, motor and emotional responses that one would experience if one were to interact with the stimulus directly (Barsalou, 2008, 2010). For example, if we see someone drinking a cup of tea, or even hear someone talking about drinking tea, this is interpreted by the brain via a partial recreation of what it would be like for us to actually drink a cup of tea ourselves. Even the sight of a teapot can trigger the firing of neurons that are actually involved in pouring and drinking a cup of tea. Crucially for this book, the embodied cognition hypothesis has important implications for our understanding of how *language* is understood and processed. Like other higher mental functions, proponents of the embodied cognition hypothesis see language as being fundamentally linked to more basic cognitive and neurobiological mechanisms (Feldman and Narayanan, 2004). Reading and hearing about actions, sensations and emotions leads us to recreate those same actions, sensations and emotions in our minds and bodies.

Evidence for the embodied cognition hypothesis and for perceptual simulation comes from large amount of empirical work on the way we understand actions, which is beginning to show that the mechanisms used in interpreting others' actions share a common representational space with mechanisms used during real action execution (Avanzini et al., 2012; Avenanti et al., 2013; Rizzolatti et al., 1996). These studies have employed a variety of research methods, including reaction time studies, neurological studies, eye tracking studies, gesture studies, questionnaires, discourse analysis and interviews. Many of these studies focus on the link between bodily action and cognition. It has been shown, for example, that parts of the frontal cortex normally associated with language processing are activated during sensorimotor action and when actions are being observed (Bonda et al., 1994). Performance on linguistic tasks has been shown to improve when participants are invited to perform corresponding physical actions (Rieser et al., 1994), words with high 'body-object interaction' ratings (such as 'scissors' and 'spoon') are recognised faster than those without (such as 'tree' and 'house') (Saikaluk et al., 2008), and the verbalisation of memories has been found to be facilitated when people assume the same body posture during recall that they had assumed when the memories were first formed (Dijkstra et al., 2007). There is increasing evidence from behavioural psychology to suggest that areas of the brain that were formerly thought to be purely sensorimotoric play important roles in language processing, and that language processing makes significant use of spatial, perceptual and visual imagery (Coslett, 1998; Coslett et al., 2002; Hauk and Pulvermüller, 2004; Rizzolatti and Buccino, 2005). On a very basic, non-linguistic level, it has been observed that when we see a person performing a particular action, such as running, gripping a pencil, laughing or crying,

the same neural motor circuits that are recruited when we ourselves perform that action are concurrently activated. Simply watching the performance of an action thus triggers corresponding motoric mental imagery. The neurons thought to be responsible for this have been referred to as 'mirror neurons' (Gallese and Goldman, 1998). Mirror neurons are thought to be partly responsible for our ability to imitate, communicate with and empathise with others (McGlone, Howard and Roberts, 2002; Stamenov, 2002).

Further support for the embodied cognition hypothesis can be found in studies showing that in word and sentence comprehension tasks, people conceptualise perceptual and motor details that go well beyond the propositions that are explicitly presented (Kok and Cienki, 2014). For example, Stanfield and Zwaan (2001) found that when participants read the sentence 'the man hammered the nail into the wall', they were quicker to recognise a picture of a nail in a horizontal orientation than a vertical one as this was compatible with the orientation of the nail in the sentence they had read. Conversely, when they read the sentence 'the man hammered the nail into the floor', they were quicker to recognise a picture of a nail in a vertical orientation. Zwaan et al. (2002) found that when people had heard the sentence 'the ranger saw the eagle in the sky' they were quicker to recognise a picture of an eagle with its wings spread, whereas when people had read the sentence 'the ranger saw the eagle in its nest', they were quicker to recognise a picture of an eagle with its wings folded. Finally, Glenberg and Kashak (2002) found that when people had been asked to read a sentence such as 'you handed Courtney the notebook', they were quicker to make a movement away from their bodies than towards them, which is consistent with the situationally appropriate movement. This evidence suggests that people construct holistic mental simulations in which they 'immerse' themselves into the scene (Zwaan, 2003).

Another set of studies that is often cited as evidence for embodied cognition are those that have shown that people use gesture to help them think. Extensive evidence exists to support the contention that gestures 'activate, manipulate, package and explore spatio-motoric information for thinking and speaking' (Kita et al., 2017: 245) and that this includes thinking and speaking with metaphor (ibid.). Kita et al. refer to this as the 'Gesture-for-Conceptualization Hypothesis'. There is substantial evidence to support this hypothesis. For example, it has been demonstrated that when people are encouraged to use gesture, they are significantly better at solving spatial problems (Chu and Kita, 2011), remembering lists of verbs (Macedonia, 2014; Macedonia and Klimesch, 2014) and performing complex statistical tasks (Rueckert et al., 2017). Finally, there is evidence to suggest that mirror neurons are involved in gesture perception, which means that mental simulation is likely to be involved in the interpretation of gestural meaning (Skipper et al., 2007). The neural networks that are involved in understanding gesture have been shown to be sensitive to contextual factors

such as the cultural background of the speaker (Molnar-Szakacs et al., 2007) and the communicative relevance of the gesture (Skipper et al., 2007). Related to this, the 'action-sentence compatibility effect' (Glenberg and Kaschak, 2002) provides evidence for the idea that language and bodily movements are related. For example, Bergen and Wheeler (2010) showed that progressive sentences about hand motion facilitated manual action in the same direction, while perfect sentences that were identical in every way except their aspect did not facilitate such action. The progressive aspect thus appears to focus attention on the nature and direction of the movement. The idea that gestures can provide insights into the embodied nature of thought is encapsulated well in Mittelberg's (2013) notion of the 'exbodied' mind.

The idea that gestures emerge from perceptual and motor simulations that underlie embodied language is particularly well developed in Hostetter and Alibali's (2008) *Gestures as Simulated Action Framework*. This is a theory of language production that attempts to account for the central role that is often played by gesture. Hostetter and Alibali argue that when people are planning to talk about an action, they simulate the action in the premotor cortex, the cerebellum and other subcortical areas, and that this simulated action activation has the potential to spread to other motor areas. When this spreading activation happens, a gesture results. Several factors are thought to affect the extent to which the activation will spread and whether the simulation will thus be realised through an overt gesture. A first factor is the extent to which the simulation involves action. Motor imagery, by definition, involves simulated action as it involves picturing the body in motion, but simulated action may also be evoked by mental imagery if one is talking about how a scene changes over time, viewing it from different perspectives, or evoking its affordances. A second factor is character viewpoint or perspective. The extent to which a speaker simulates actions using motor imagery or animated visual imagery is likely to vary according to whether the speaker is describing his or her own actions. Hostetter and Alibali (2010) found that participants gestured at a higher rate when they were describing a pattern that they had actually made than when describing a pattern they had only viewed. A third factor is individual differences. Individual differences have been observed in people's propensity to gesture: speakers who typically rely on simulations of perception and action are particularly likely to use gesture when speaking. Hostetter and Alibali argue that these individual differences may depend on the strength of the neural connections between the premotor and motor areas, and that the strength of these connections can result from genetic factors, experience or individual differences in the functioning of the nervous system. In support of this claim they cite their finding that speakers with stronger spatial skills tend to gesture more than speakers with weaker spatial skills (Hostetter and Alibali, 2007). A fourth factor is the communicative setting. Studies have shown that teachers increase their use of gesture when teaching particularly difficult

concepts or when they believe that their interlocutor may have difficulty understanding them (Alibali and Nathan, 2007). Finally, the nature of the material being discussed and the speaker's attitude towards that material are also likely to affect the use of gesture. McNeill (2005) has shown that speakers are particularly likely to gesture when expressing contrast or when conveying information that is more newsworthy or crucial to the ongoing conversation in general, a feature that he labels *communicative dynamism*. To sum up, the fact that meaningful gestures often accompany speech and that gesture use varies according to different communicative contexts provides compelling support for the idea that cognition is often embodied.

Indirect evidence for the embodied nature of thought can also be found in everyday language. For example, canonical word orderings in English reflect the ways in which we interact with the world through our bodies. As we have eyes in our head, not our feet, look forward rather than backward and stand upright, elements of language that are in front, above and vertical tend to be foregrounded and are mentioned first. For these reasons, it is more conventional to talk about things being 'high and low', 'above and below' and 'front and back' rather than 'low and high', 'below and above' and 'back and front' (Benor and Levy, 2006). Thus words to do with being upwards and forwards are more likely to precede their counterparts in canonical expressions. Using the Web IT 5-gram corpus as his data source, Louwerse (2008) calculated the raw and relative frequencies of iconic and non-iconic word orderings of seventy-one word pairs that have been widely studied in the literature, and found that they were significantly more likely to occur in their canonical (more 'embodied') order than in their non-canonical (less 'embodied') order.

Support for the embodied cognition hypothesis can also be found in the results of eye tracking studies. Huette et al. (2014) had participants look at a blank screen while listening to stories that contained either past progressive sentences or simple past sentences. They found that participants who heard past progressive sentences moved their eyes around the screen significantly more than participants who heard simple past sentences, who fixated on a single spot. This finding is in line with the fact that the past progressive focuses attention on the unfolding aspects of an event, whilst the simple past focuses attention on its end point. Their findings thus provide support for the idea that grammatical aspect alters the way in which events are construed. Huette et al. conclude that 'there is nothing passive about passive listening' and that 'the eyes are actively moving in a way that reflects subtle grammatical differences in the linguistic input' (ibid., 7). Similarly, other studies have found that the speed with which people move their eyes towards a particular target varies according to the amount of speed that is encoded in the verb that they hear. For example, people have been found to move their eyes more quickly towards a target after having heard sentences such as 'the lion dashed toward the balloon'

than after having heard sentences such as 'the lion ambled toward the balloon' (Lindsay et al., 2013; Speed and Vigliocco, 2014).

An important feature of the embodied cognition hypothesis is that it emphasises the role played by emotion in cognitive processing. Emotional involvement has been shown to be an integral, indispensable part of much so-called logical or abstract reasoning. It facilitates cognition, supports understanding and enhances long term memory (Storbeck and Clore, 2008). Evidence for this comes from a number of empirical studies. The comprehension of sentences describing emotional situations (Havas et al., 2007) has been found to be affected by the listener's bodily state, and it has been shown that when people engage emotionally with a piece of information, they are more likely to understand it and remember it (Webb, Miles and Sheeran, 2012). In one's first language, words with high emotional content are better retained than more neutral ones (Kensinger and Corkin, 2003). Based on his study of patients who had both decision-making defects and emotional disorders, Damasio (2006) advanced what he refers to as the 'somatic marker hypothesis'. 'Somatic markers' are physical responses that are associated with emotions, such as the association of rapid heartbeat with anxiety or of nausea with disgust. He found that in some of his patients, flawed reasoning, emotional impairment and a reduction in the activation of somatic markers stood out together as consequences of a specific brain lesion. He concludes from this that emotion plays a key role, albeit a subconscious one, in what he refers to as 'the loop of reason' (ibid., xvii), arguing that increased levels of emotion increase the saliency of a premise and in so doing, bias decision-making in favour of that premise. He also argues that heightened emotional engagement helps us to hold in mind multiple facts that must be borne in mind during a decision-making process (xviii). For Damasio, emotion also plays a key role in intuition, which he describes as the speedy cognitive process through which we reach a conclusion without being aware of all the logical steps. To use his words, intuition is 'rapid cognition with the required knowledge partially swept under the carpet' (ibid., xix). A full account of embodied cognition must therefore include consideration of the way in which we interact emotionally with our surroundings as well as physically. Although embodied cognition cannot account for all aspects of cognitive processing (Goldinger et al., 2016), it does explain a number of features of human cognitive and communicative behaviour, which, as we will see later, include the ways in which we understand and use metaphor.

Embodied cognition is not always a conscious process. Lakoff and Johnson (1999) propose three levels at which embodied cognition operates: *neural embodiment* involves the neurological architecture that underpins embodied thinking; the *cognitive unconscious* concerns the rapid underlying mental operations that we use automatically, without being aware of them; and the *phenomenological level* concerns our awareness of our bodies, mental states,

our environment and our physical and social interactions. Gibbs (2005) makes the important point that in order to acquire a full understanding of the nature of embodied cognition, we need to focus on all three levels of embodiment, and most importantly, the interactions between them.

In response to the growing number of theories that have been proposed pertaining to embodied cognition, Meteyard et al. (2012) conducted an extensive review of the literature and concluded that there is no empirical support for either strongly embodied theories of cognition or for completely disembodied theories of cognition. They suggest that the remaining theories of embodied cognition are best viewed as sitting along a continuum. This goes from 'secondary embodiment', where semantic content is independent but associated with sensorimotor systems, through 'weak embodiment', where semantic content has partial dependence on the sensorimotor system, to 'strong embodiment', which asserts that there is complete dependence on the sensorimotor system. One of the aims of this book is to explore the ways in which people's experiences of embodied metaphor position themselves along this continuum and then identify factors that contribute to this positioning.

Embodied Cognition and the '4E' Approach

Human cognition is not confined to the individual. Rather, it is shaped by one's physical and social environment as well as one's personal and social history. It is thus environmentally, socially and temporally 'distributed' (e.g., Jensen, 2013; Semin and Cacioppo, 2009). The notion of distributed embodied cognition emphasises the fact that we operate in a world populated by other people and things, and that this world changes over time.

The '4E' approach to cognition emphasises the fact that as well as being ***e****mbodied*, cognition is also ***e****xtended*, ***e****nactive* and ***e****mbedded*. The term 'extended' emphasises the relationship between internal embodied cognition and the physical and social environment. It takes as its starting point the fact that we understand things in terms of what they mean to us and how we can best make use of them or interact with them (Glenberg, 1999; Glenberg and Kaschak, 2002) and what it is that they afford us (Gibson, 1977). The term 'enactive' emphasises the active nature of thinking and foregrounds the role played by gesture and other types of bodily movement in the thinking and communication. The term 'embedded' emphasises the fact that all cognition takes place in a context and that contextual features very much shape the way we think. There is considerable overlap between the 4Es but taken together, they emphasise the fact that both language and cognition are:

> materially embodied, culturally/ecologically embedded, naturalistically grounded, affect-based, dialogically coordinated, and socially enacted. (Thibault, 2011: 211)

As we will see later in the chapter, the 4E approach also allows us to see metaphor as a more dynamic phenomenon than traditional embodied approaches. This is the position that I adopt in this book.

1.3 Embodied Metaphor

In recent years, discussions of embodied cognition have been extended to metaphor, in particular to those metaphors that relate abstract content to more concrete referents, such as, for example, AFFECTION IS WARMTH, DIFFICULTY IS WEIGHT or IMPORTANCE IS SIZE.[2] Metaphors such as these are often referred to as 'primary metaphors' (Grady, 1997a). Primary metaphors are metaphorical correspondences or 'mappings', for which there is thought to be an experiential basis as they are derived from real-world experiences, so the three metaphors that I have just mentioned can be accounted for by the facts that: when we think about someone we are close to it can make us feel warm inside; heavy objects are harder to carry; and important things and people tend to be bigger than us, at least when we are children. Primary metaphors are thought to develop in early childhood as a result of correlations between different yet related experiences, and are therefore sometimes referred to as 'correlational' metaphors. For example, as infants, when we experience affection from our parents, it usually takes the form of closeness and warmth, and as adults, we continue to describe affection metaphorically, describing our relationships as 'close' and 'warm', leading to the AFFECTION IS WARMTH metaphor that has just been mentioned. Another primary metaphor is the TIME AS MOTION ALONG A PATH metaphor (Ibid., 294). When we walk along a path, the section of the path where we have already walked is behind us and the section of the path where we have yet to walk is in front of us. We thus perceive of the past as being metaphorically 'behind' us and the future as being metaphorically in front of us. We can see from these examples that there is no clear-cut distinction between 'the body' and 'the mind' and that physical actions influence abstract thought. It is thus argued that abstract concepts originate in our experiences of our bodies as they interact with the physical world, and these experiences constitute our primary, most basic source of understanding (Bergen, 2012).

Relationships that we perceive in our environment are internalised, and this is thought to lead to the formation of 'cross-domain mappings' (Lakoff and Johnson, 1980). These mappings are formed between 'source domains' and 'target domains'. Source domains are concrete concepts that we experience through sensorimotor interactions with the world. In the aforementioned examples, these would be WARMTH, WEIGHT and SIZE. Abstract domains are the less-concrete concepts that are being considered or discussed. In the aforementioned examples, these would be AFFECTION, DIFFICULTY and IMPORTANCE. The theory of neural metaphor (Lakoff, 2008, 2012) goes one