# *Wheels Within Wheels:* A Causal Treatment of Image Schemas in An Embodied Storytelling System

Philipp WICKE [a] [1] and Tony VEALE [a]
[a] *University College Dublin, Belfield, Dublin 4, Ireland*

**Abstract.** Telling a story requires the internal interrogation and reassembly of our memories of situated actions in past events. Expressing those actions in a context of embodied story-telling often allows for, or even demands, the use of physical gestures. We believe that such gestures are most effective when they make use of the recurrent spatial patterns of human cognition, image schemata. This extended abstract outlines a novel approach to grounding a computational storytelling system in an embodied robotic agent that is capable of making such gestures, and argues for why image schemata make the ideal glue for linking the causal structures of plot generation to the gestures of bodily expressiveness.

**Keywords.** image schema, embodied cognition, computational creativity, storytelling

## 1. Introduction

It has been argued that image schemata are a fundamental basis of how humans express themselves, whether speaking of concrete or abstract ideas. Image schemata are recurring cognitive structures that are shaped by our bodily interaction with our physical environment. The definition has been most prominently championed, among others, by Johnson [1] and Lakoff [2]. These definitions presuppose and apply the theory of embodied cognition, in which cognitive processes are to be understood as processes of the mind and the brain, including the body and its environment, challenging the doctrine of pure Computationalism [3]. As image schemas impose meaningful structures on our bodily movements through space and time, they can be observed not just in our spoken language but in our body language too, i.e. gestures [4][5][6].

This extended abstract proposes an approach to utilizing image schemas in an embodied story-generation and story-telling system. As such, the domains of Cognitive Linguistics/Semantics and Computational Creativity overlap in a framework that aims to lift the creative processes of stories *out of their container*, in this case an embodied NAO robot [2]. Building on our interactive storytelling framework [7][8] we investigate the use

[1]Corresponding Author: Philipp Wicke, University College Dublin, Belfield, Dublin 4, Ireland; E-mail: philipp.wicke@ucdconnect.ie
[2]For more information see www.softbankrobotics.com/emea/en/nao

of image schemas as a cognitive framework for a storytelling robot. For this investigation we have correlated a set of 9 prominent image schemas with more than 800 story actions (plot verbs) from the storytelling system. Prior to the empirical evaluation of gestures based on the image schemas incorporated into the system, this paper presents a preliminary study which examines the database that has been created to combine gestures, schemas and story-actions. This examination lays the groundwork for a consideration of future work and expected progress on the proposed framework.

## 2. Image Schema for Robotic agents

Our goal is to build an engaging robotic story-teller than exploits its physical presence to lend embodiment to the symbolic narratives of an automated story-telling system. We consider here the use of image schemas to enrich the representation of story actions and their corresponding gestures in earlier versions of the work [7][8]. By giving the system an image-schematic understanding of its own actions and gestures, we hypothesize that it will be better positioned to choose apt gestures in context, or to mold new stories around these gestures (as opposed to molding gestures around stories). Image schemas are powerful abstractions that capture the spatial/temporal logic of an action, as shown in the following instantiations of the *outward motion*:

> John went out of the room. Pump out the air. Let out your anger. Pick out the best theory. Drown out the music.
> *(Selected examples from Johnson [1] p. 32 - Figure 4)*

We choose our core set of image schemas to suit the gestural capabilities of our robot, a humanoid Nao robot from the French company Aldebaran Robotics (recently acquired by Softbank Robotics). This bipedal and bimanual robot, released in 2008, has 25 degrees of freedom, and can move its arms in ways that accommodate many human gestures. We do not need very many image schemas to express a wide array of different actions, so we choose only those that match the capabilities of the robot. Consequently, we propose the selection of schemas outlined in Table 1.

| Image Schema | (Direction) Gesture | (Inverse Direction) Gesture |
|---|---|---|
| UP-DOWN | (UP) Both arms up | (DOWN) Both arms/torso lower |
| FRONT-BACK | (FRONT) Both arms stretching to front | (BACK) both arms moving back |
| CENTRE-PERIPHERY | (CENTER) Hands unite from sides | (PERIP.) Hands depart from centre |
| CONTACT | (SAME) Both hands touch in front | (SAME) Both hands touch in front |
| IN-OUT | (IN) arm-1 encloses, arm-2 reaches in | (OUT) arm-1 encloses, arm-2 reaches out |
| SURFACE | (SAME) Both hands circle horizontally | (SAME) Both hands circle horizontally |
| CYCLE | (SAME) Both hands circle vertically | (SAME) Both hands circle vertically |
| S-P-G | (SAME) arm-1 moves vertically 3 times | (SAME) arm-1 moves vertically 3 times |
| NEAR-FAR | (NEAR) arm-1 fixed front, arm-2 unites | (FAR) arm-1 fixed front, arm-2 departs |

**Table 1.** Our 9 chosen image schemas and their corresponding gestures. If a schema requires a direction along an axis, two gestures are provided, otherwise the same gesture is given in both columns. S-P-G is the Source-Path-Goal image schema.

Table 1 provides an overview of nine image schemas. Of these, five are sensitive to direction, meaning that a corresponding inverse gesture along the opposite direction also exists. For example, for the UP-DOWN schema both UP and DOWN have their own distinct gestures, though each can be seen as a directional inflection of the other. In all then, the 9 schemas give rise to 14 distinct gestures. The association between gesture and image schema builds on the evidence provided in other studies [4][5][6] and its validity for our framework will be the subject of further investigation. See Figure 2 for an overview of the proposed improved framework. To connect these image schemas to the storytelling system, we next consider the database of actions on which the storytelling system is built.

## 3. Image Schemas in Creative Story-Generation and Story-telling

We tie these schemas and their gestures to specific actions in the repertoire of the story-generation system. We build upon the *Scéalextric* system of [9], since its approach is explicitly action-centred. Each *Scéalextric* story is composed by chaining together actions (chosen from a pool of 800 plot verbs) into a causal path (explanation of this causal chaining can be found in [9]). Actions connect to subsequent actions via causal connectives, such as *so*, *then* and *but*. A key part of *Scéalextric* is its large set of causal tuples, which indicate how each of its 800 actions might result in another. In our previous work [8], we tied the movements of the NAO gesture library to the actions of the *Scéalextric* causal graph, so that a path through the latter could be physically embellished (or, in some cases, pantomimed) by the robot. This mapping of gestures to actions was done in the absence of any mediating abstraction, and was thus opportunistically ad-hoc. However, we can now use image schemas as the motivating abstraction to tie gestures to story actions. Given the causal underpinnings of *Scéalextric*, this will subsequently allow us to look for recurring causal patterns between the image schemas themselves.

*Scéalextric* uses the causal connectives *but*, *but later*, *yet*, *until*, *and*, *then*, *because*, *though* and *so* to connect two successive story actions. It is self-evident that *but* and *but later* form a group of connections (hereafter the BUT-Group) reflecting how a successor action happens contrary to its predecessor, e.g. "*Faust **fell in love** with Gretchen, but later he **betrayed** her.*". In the group of connections *then*, *because*, *though* and *so* (hereafter the SO-Group) we expect a strong causal continuation, e.g. "*Mephisto **stood up for** Faust, so Faust **was impressed by** Mephisto.*". In the last group of connectives *yet*, *until* and *and* (hereafter the AND-Group) we neither expect nor are surprised by a given continuation, e.g. "*God **challenged** Mephisto the Devil and Mephisto the Devil **enchanted** Faust.*". We can now analyze our mappings of image schemas to actions using the following hypothesis:

> The continuity of story actions is reflected in the interconnection of image schemas over the course of a story.

## 4. Analysis of Database

The first analysis compares how many image schemas follow the continuation of the actions from the storytelling system. For that, we investigate the continuation of directionality of the image schema through the course of a story, i.e. the UP schema repeatedly following schemas related to direction away from the robot (FAR, PERIPHERY). We expect this continuation of direction in stories of the SO-Group. The storytelling system draws the causal connectives from a database, which connects the 800+ actions in about 3700+ combinations with causal connectives, e.g. ”trusted by BUT stalked by”. The first result divides the connectives into the aforementioned three groups and tests if the continuation holds for the image schema, e.g.: ”*Faust* ***fell in love*** *with Gretchen, but later he* ***betrayed*** *her.*”. Here, **fell in love** is mapped onto the schema ”IN”, because falling in love is the conceptual metaphor for the example of entering the abstract image schema *CONTAINMENT* [2]. **betrayed** is mapped onto the schema ”BACK” (following the metaphorical stab in the back, or to be attacked in the back). Importantly, this first investigation simply looks for the continuation of the image scheme irrespective of the direction (e.g., BACK-FRONT treated as one image schema). The difference between the BUT-Group and the SO-Group is 5.84%, i.e. in the SO-Group just 5.84% more actions show causal continuity in their use of the same image schemas.

In Figure 1, we see the relative frequency of each image schema pairs connected by SO-Group connectives (left) or the BUT-Group connectives (right). Each row depicts for each image schema the percentage of occurrence for the following image schema.

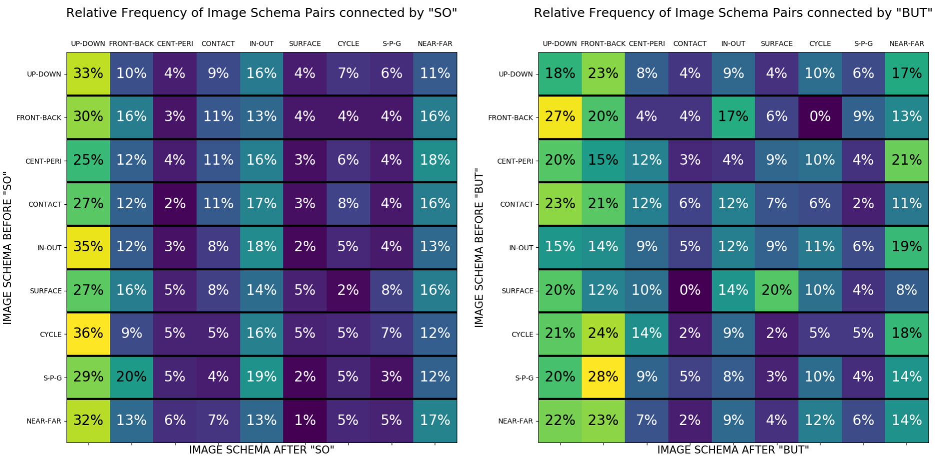


**Figure 1.** Relative Frequency of Image Schema Pairs connected by ”SO” (left) and ”BUT” (right).

In the second analysis, we investigate the image schema with respect to their direction. We group the image schemas according to their spatial relation, i.e. the first group (Group A) comprises image schemas related to movement towards the robot: BACK, CENTER, IN, NEAR. The second group (Group B) comprises image schemas related to movement away from the robot: FRONT, PERIPHERY, OUT, and FAR. The last group (Group C) comprises all other image schemas: UP, DOWN, CONTACT, SURFACE, CYCLE, and SOURCE-PATH-GOAL. This second investigation allows us to tighten the association between spatial domain and image schema. The difference between the BUT-Group and the SO-Group is 7.84%, i.e. in the SO-Group 7.84% more actions are continuous with regard to their image schema group. While this is a marginal difference, it indicates that this preliminary trend does not confirm but also does not contradict our hypothesis.

## 5. Conclusion and Future Work

This research presents a novel way of investigating image schemas in the realm of computational embodied storytelling, and sketches a framework that utilizes a storytelling system to explore the causal connections between image schemas. While initial results are minimal at best, these investigations point towards easily extensible further studies and improvements. The current mapping between image schemas and actions is a definite weak point and potential explanation for the weak results. In future research, we plan to refine the mapping between the image schemas and actions in order to test the proposed hypothesis in general terms. In Figure 2 we can observe the two crucial links between story actions and image schemas and image schemas and gestures. Necessary evidence in previous research has been discussed to justify this investigation. Consequently, we want to investigate the benefits of grounding actions and gestures in a common set of image schemas in contrast to the previous ad-hoc mappings. This common set of image schemas will draw from the idea of *Conceptual Scaffolding* [11] which provides a simple calculus for reasoning about image schema. This approach will allow us to define a set of gestural expressions of those inherently spatial schemas. The implementation of those schemas will be guided by recent approaches of [12], who use image schema for conceptual blending and with that provide a formalized logic language for image schema. Strengthening the links of this framework will help to provide the necessary data which will allow a proper empirical evaluation.

## References


[1] M. Johnson *The body in the mind: The bodily basis of meaning, imagination, and reason* University of Chicago Press, (2013)

[2] G. Lakoff *Women, fire, and dangerous things* University of Chicago press, (2008)

[3] A. Chemero *Radical Embodied Cognitive Science* MIT Press, (2011)

[4] A. Cienski *From perception to meaning: Image schemas in cognitive linguistics - Image schema and gesture* De Gruyter, *421:442*, (2005)

[5] I. Mittelberg *Gestures as image schemas and force gestalts: A dynamic systems approach augmented with motion-capture data analyses* De Gruyter, *Vol. 11 - Cognitive Semiotics* (2018)

[6] B. Ravenet, C. Clavel and C. Pelachaud *Automatic Nonverbal Behavior Generation from Image Schemas* Proceedings of the International Conference on Autonomous Agents and Multi-agent Systems, (2008)

[7] P. Wicke and T. Veale *Interview with the Robot: Question-Guided Collaboration in a Storytelling System* Proceedings of the International Conference on Computational Creativity, (2018)

[8] P. Wicke and T. Veale *Storytelling by a Show of Hands: A framework for interactive embodied storytelling in robotic agents* Proceedings of the Conference on Artificial Intelligence and Simulated Behavior, (2018)

[9] T. Veale *A Rap on the Knuckles and a Twist in the Tale From Tweeting Affective Metaphors to Generating Stories with a Moral.* Proceedings 2016 AAAI Spring Symposium Series, (2016)

[10] T. Veale *Déjà Vu All Over Again.* Proceedings of the International Conference on Computational Creativity, (2017)

[11] T. Veale and M. Keane *Conceptual Scaffolding: A spatially founded meaning representation for metaphor comprehension.* Computational Intelligence 8.3 *494-519*, (1992)

[12] M.M. Hedblom, O. Kutz and F. Neuhaus *Image schemas in computational conceptual blending.* Cognitive Systems Research 39 *42-57* (2016)

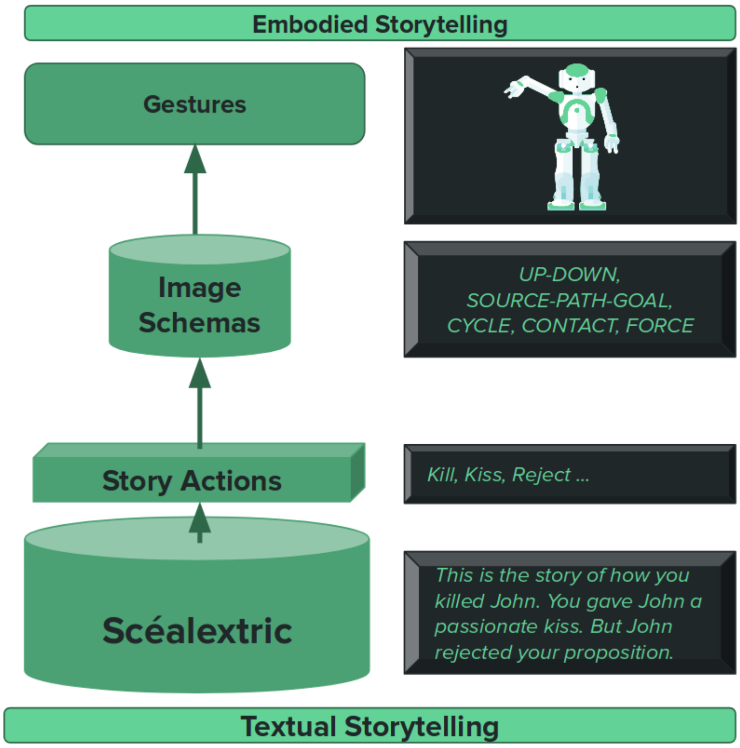


**Figure 2.** The figure depicts the proposed framework, which starts from the textual storytelling utilizing the automated storytelling system *Scéalextric* to create stories. From those stories, actions can be retrieved which will be mapped onto a set of image schemas. Subsequently, those image schemas can be related to gestures to achieve an embodied storytelling approach.