


Relations Among MOCAP and Textual Data of Motion Verbs: A Distance Calculation Perspective

Marietta Sionti, Athena Research and Innovation Center in Information Communication and Knowledge Technologies, Greece & Neurocognition and Action Group, Cognitive Interaction Technology Institute, Bielefeld University, Germany

Panagiotis Kouris, Athena Research and Innovation Center in Information Communication and Knowledge Technologies, Greece

Chrysovalantis Korfitis, Athena Research and Innovation Center in Information Communication and Knowledge Technologies, Greece

 <https://orcid.org/0000-0002-1697-5537>

Vasiliki Moutzouri, Athena Research and Innovation Center in Information Communication and Knowledge Technologies, Greece

Stella Markantonatou, Athena Research and Innovation Center in Information Communication and Knowledge Technologies, Greece

ABSTRACT

This is a study on language and action that tries to shed light on their conceptual correspondence in terms of embodiment. The linguistic phenomenon of lexical aspect/Aktionsart is studied in connection to joint angles and time. For the purposes of this research, data concerning the usage of a set of Modern Greek verbs were collected and annotated; in addition, motion data were captured for the same verbs. First, the distance between actions and verbs was calculated revealing a strong connection between certain types of language data, lower limb motion, and time: an extended use of the lower limbs is related to longer or repetitive actions while time expresses duration.

KEYWORDS

Abstract Meaning Representation, Cosine Distance, Earth Mover's Distance, Embodiment, Euclidean Distance, Human-Robot Interaction, Intangible Cultural Goods, Kinematics, Lexical Aspect/Aktionsart

INTRODUCTION

According to the UNESCO Convention on the Safeguarding of Intangible Cultural Heritage and the Greek Law 3028/2002 "On the Protection of Antiquities and Cultural Heritage in general" (Official Government Gazette 15/A' / 28.6.2002) as intangible cultural goods are considered different facets of everyday language, such as verbal expressions, oral traditions, myths, songs, etc. (ayla culture, 2020). Our goal is to analyze language as a cultural and cognitive phenomenon related to action, under the prism of embodiment. For this reason, we attempt to ground the abstract linguistic notion of Aktionsart/lexical aspect to concrete low level kinematic features, captured with a Synertial biomechanics suit. Since the majority of relevant literature is deploying visual data, the proposed linkage is the first step of an innovative analysis with potentially far-reaching effects, for instance, the

DOI: 10.4018/IJACDT.2020070104

results could be embedded in robots in order to enhance embodied human-robot interaction, human activity recognition, event segmentation and labeling of sensorimotor data.

To this end, first we captured kinematic data about the actions denoted by a set of motion verbs of Modern Greek, namely actions of pushing, pulling, hitting, and beating. Next, we collected 150 sentences for each of the studied verbs from the web and annotated them drawing on the *Abstract Meaning Representation* framework (AMR); (Donatelli, Schneider, Croft, & Regan, 2019). Eventually, we propose an extension of the AMR schema suitable for the Modern Greek aspectual system. Next, for the same actions we obtained mocap and linguistic annotation data and calculated the distance between them. Finally, we combined kinematic and textual results, in order to ground aspectual features (such as actor, telicity, adverb for aspect declaration, etc.) to specific movements for the needs of human robot interaction. In the preliminary results, the feature “*time and lower limbs*” that occurs in the mocap data returns the best intersection of sets of close verb pairs obtained from the kinematics and linguistic data, suggesting that the particular lexical aspect features obtained from corpus data may be related with the internal temporal constitution of these actions.

Aspect is a core linguistic property of structures with verbal features (verbs and deverbal nominals); it is about temporal and frequency information that results from the semantics of the verbal element and the overall structure that is headed by it. Comrie (1981, p. 3) supports that aspect provides different ways of viewing the internal temporal constituency of the situation. Here we are not interested in how the aspect of a sentence is calculated (for instance, see Dowty’s Montagovian approach (1979)) but in the verbal lexical aspect or Aktionsart. Verbal lexical aspect has to do with the temporal structure of a verb’s meaning and is independent of the various structures that the verb may head. Venerated approaches to the issue are those of Zeno Vendler (1957) and Comrie (1981) while extensive work on Modern Greek can be found in Moser (2013).

This work is about verbs denoting movement; motion events, by their nature, have an inherent temporal dimension and a certain duration, which is considered to play significant role in the mental simulation of an action both in the execution of the movement per se and the linguistic expression of the real world actions (Bergen & Chang, 2005; Matlock et al. 2005; Zwaan, 1999; Barsalou, 2009), that have been previously observed and learnt by mirror neurons (Fadiga et al., 2006; Arbib, 2008).

Recently, human motion analysis and event segmentation have attracted the interest of a range of researchers, namely computer vision engineers, computational linguists and psychologists: engineers attempt to match language with videos in order to assist activity recognizers while linguists and psychologists anticipate that the visual data will offer crucial information to linguistic analysis and embodiment. The driving force for this extensive engagement is useful and promising applications. Thus, high level applications, such as security surveillance environments, healthcare systems, human performance analysis in sports activities, human-computer interfaces, educational systems, etc., require low level technology, e.g. feature extraction and representation, mainly based on single person activity recognition (Ke et al. 2013; Yu, Teo, Yang, Fermueller & Aloimonos, 2013). These features may vary depending on the approach. In the case of the generic model recovery a 3D model is required, while the motion-based model exclusively utilizes motion characteristics. The appearance-based model is a 2D shape model derived from images and videos (Ribeiro, Santos-Victor & Lisboa, 2005). In this line of research, the majority of researchers who attempt to detect activities work with three dominant motion features, namely space, time and frequency (Ke et al. 2013). Motion verbs, in particular, have received some considerable attention in studies that try to ground language in sensorimotor data because they are among the linguistic entities that are thought to be considerably grounded in action (for visual data see Pastra & Aloimonos (2012), for kinematic data see Sionti, Claudino, Aloimonos, Rose & Markantonatou (2011)).

The paper is structured as follows: A brief review of the (vast) literature on kinematics and aspect is given first. Next, we discuss the suitability of kinematics for an early stage investigation of aspectual elements and we describe the mocap collection process. We analyze our methodology for

16 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/article/relations-among-mocap-and-textual-data-of-motion-verbs/259879

Related Content

Urban Interiors and Interiorities

Suzie Attiwill (2020). *Cultural, Theoretical, and Innovative Approaches to Contemporary Interior Design* (pp. 58-67).

www.irma-international.org/chapter/urban-interiors-and-interiorities/249450

Mixed Reality Boundaries in Museum Preservation Areas

Bernardo Uribe, Luis Miguel Méndez, Andrés Tovar, Jean Pierre Charalambos, Olmedo Arcilaand Álvaro David López (2013). *International Journal of Art, Culture and Design Technologies* (pp. 63-74).

www.irma-international.org/article/mixed-reality-boundaries-in-museum-preservation-areas/90161

Transforming Culture to Stimulate Economic Development: Lessons from Singapore

Balbir Bhasinand Lee Keng Ng (2016). *International Journal of Art, Culture and Design Technologies* (pp. 48-58).

www.irma-international.org/article/transforming-culture-to-stimulate-economic-development/163118

Identity Transformation through Collaboration: Narratives of 'Becoming an Architect'

James Thompson (2017). *Collaboration and Student Engagement in Design Education* (pp. 332-353).

www.irma-international.org/chapter/identity-transformation-through-collaboration/165689

Tap and Swipe Activity Zones on Smartphones for Novice Older Adults

Paula Alexandra Silva, Roxanne Leitãoand Maureen K. Kerwin (2018). *International Journal of Creative Interfaces and Computer Graphics* (pp. 18-39).

www.irma-international.org/article/tap-and-swipe-activity-zones-on-smartphones-for-novice-older-adults/218885