Comparing linguistic classifications with sensorimotor data of English and Greek verbs of motion

Marietta Sionti[1], Leonardo Claudino[2], Carolyn P. Rose[3], Stella Markantonatou[4], Yiannis Aloimonos[2]

[1]University of Athens, [2]University of Maryland, [3]Carnegie Mellon University, [4]Institute for Speech and Language/ 'R.C.Athena'

Abstract: We combine linguistic knowledge from corpus data with sensorimotor data obtained experimentally in an effort to better specify the minimum conceptual representation of a motion event that distinguishes it from all other events. Sensorimotor data are collected by measuring the performance of speakers of Modern Greek and American English. We focus on the clustering of motor actions and its correspondence to previous linguistic classifications of both languages.

1 Introduction

We combine corpus driven linguistic knowledge with experimentally obtained sensorimotor data in an effort to better specify the minimum conceptual representation of a motion event that distinguishes it from all other events. We use American English and Modern Greek data as a case study, in order to enable crosslinguistic analysis. We hope that our work will contribute to a better understanding of both language and perception. Sensorimotor data, here collected by measuring the motor behavior of speakers of American English and Modern Greek, allow for linking embodied experience and language (image schemas are learnt as a sequence of interrelated sensorimotor patterns (Lakoff & Johnson, 1999)). We expect that by delineating conceptual representations of motion events in this way, we will be able to (i) better understand semantic classifications of verbal predicates (ii) perceptually ground abstract notions, eg *transitivity*, that have traditionally been used in linguistics to study and classify verbal semantic and syntactic properties, and by combining them with corpus driven data, (iii) offer a quantitative answer to long standing syntactic questions such as the distinction between "argument" vs. "adjunct", whose binary nature has been questioned (Galen, Grenager&Manning, 2004).

This paper focuses on the collection of sensorimotor data, the clustering of

motor actions and its correspondence to previous linguistic classifications (Levin, 1993; FrameNet; Antonopoulou, 1987). The detailed sensor data are analyzed to identify latent factors that represent stable patterns across the many dimensions of low level data. These factors appear as discrete sets (*synergies*) of joint angles and orientations associated with each action and are correlated with linguistic descriptions.

We draw on the extensive prior work related separately to Cognitive Science (Jackendoff, 1990; Feldman, 2006), Mirror Neurons and their impact on language (Fadiga *et al*, 2006; Arbib, 2008; Kemmerer, 2006) and Computer Science (Santello, 1998; Troje, 2002).

2 Linguistic Classifications

2.1 Modern Greek:

There exist two studies on the classification of Modern Greek Motion Verbs (MGMV) (Antonopoulou, 1987; Mpasea, 1996). MGMV exhibit an overall semantic structure found with motion verbs of several Indo-European languages and, at the same time, present some certain aspectual idiosyncrasies. Antonopoulou (1987) adopted prototype theory as the most suitable method for the investigation for MGMV; prototype theory is by default closer to the cognitive approach adopted here.

Antonopoulou's taxonomic sets were defined with the use of two groups of criteria:

- Criteria of the first group: transitivity, causativity, agentivity, intentionality and aspect; none of them can be measured with sensorimotor methods at the moment.
- Criteria of the second group: change-of-location, directionality, path, dependent motion, change-of-orientation, manner, medium and instrumentality.

2.2 American English

Though the difference in perspective of Levin's (1993) English Verb Classes and Framenet's categorization is well attested (Baker & Ruppenhofer), both these classifications are important for this work. Levin's classes are based on semantic grouping and valence alternations. Very much like Antonopoulou, Levin offers a rich anthology of verbs enriched with syntactic information that is crucial for our long standing goal, namely the distinction *argument* vs *adjunct*. On the other hand, Framenet's grouping of words according to conceptual structures can be easily matched to Antonopoulou's second group of criteria and, finally, to sensorimotor data.

3 Sensorimotor experiment

3.1 Verb collection criteria

The verbs used for the experiments fulfilled the following criteria: (i) one human participant per action, including both intransitive and transitive verbs, and (ii) keeping to more literal meanings mainly due to lab limitations (Table1).

3.2 Experimental procedure:

Method: The detailed sensor data were analyzed to identify latent factors that represent stable patterns across the many dimensions of low level data. These factors appear as discrete sets (*synergies*) of joint angles and orientations associated with each action.

Equipment: A full body Moven system contains 16 inertial motion trackers. Each sensor module comprises 3D gyroscopes, 3D accelerometers and 3D magnetometers. Using advanced sensor fusion algorithms (Moven Fusion Engine) the inertial motion trackers give absolute orientation values which are used to transform the 3D linear accelerations to global coordinates which in turn give the translation of the body segments. The advanced articulated body model (23 segments and 22 joints biomechanical model) implements joint constraints to eliminate any integration drift or foot-sliding.

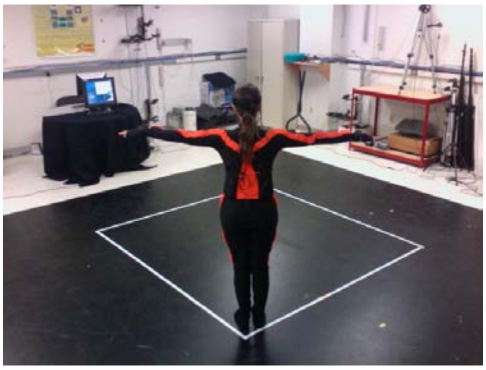

fig.1 A subject while performing an action

Subjects: The age range of the subjects was 25-30 years old and their sexual distribution 5 men and 3 women. Being native speakers, they were encouraged to implement each meaning according to their intuition.

Phases: The sensorimotor experiment was divided in two phases; (i) capturing of the action performance- this part yielded the main dataset for Greek and English (8 subjects each) and (ii) correspondence between the languages.

Action performance: This was a step-wise procedure. The verb or the sentence was uttered by the experimenter. When the verb was performed with the body of the subject only, action was limited to a floor area restricted by a quadrangle. In order to normalize the distance, subjects were encouraged to start acting at a specific corner of the quadrangle (fig.1).

Although several actions could be implemented by using only the body and the prerequisite was to involve as few objects as possible, the subjects asked for items that could be found in the lab:

- a step (verbs 8, 10, 15)
- a ramp (9, 11)
- one or several balls (5, 14)
- table, book, cylinder, chair (22-25)
- chair (20, 21)

In order to standardize the procedure, the same objects were used throughout the experiments (whenever an object was required).

Correspondence of the verbs in the two languages: The Greek participants performed the Greek verbs of motion and one of them was videotaped. Ten English speakers were shown the video segments and were asked for the corresponding English verb that would best describe each action. In controversial answers, we substituted the open question with multiple choices, complementing them with similar entries from WordNet. In the cases that the problem persisted, we asked the English participants of the sensorimotor experiment to perform both choices. Therefore, two tendencies were observed: (a) 1:1 correspondence between the verbs of the two languages and overlapping in the meanings and (b) participants not feeling confident both about the meaning of the verb and how to perform the corresponding action.

4 Analysis of the data and results

4.1 Analysis

In order to describe the motor representation of each verb, we extract its average action. These average actions are normalized in length and further "stacked", forming the base motor data matrix of our work. That matrix is called the Principal Component Analysis (PCA) and is processed as in (Santello, 1998), allowing a two-dimensional visualization of the action scatter.

The two main visual groups of actions are (i) the rectangular that includes walking like actions (leg-related), and (ii) the blue ellipsis that includes manipulation of an object (arm-related) (fig.4, 5). They are projected on both PC1 (distinct use of hips) and PC2 (emphases on knees and shoulders) (fig.6, 7). Therefore, all these actions have approximately the same profile in terms of joints-angles.

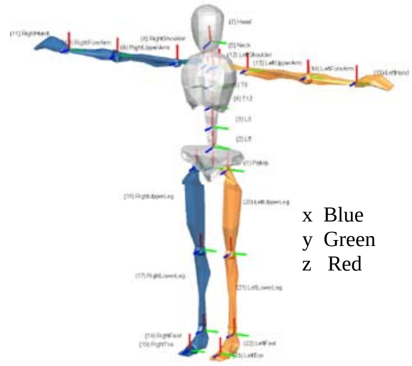

fig.2 X Moven Suit's axes

Principal Component 1 (PC1, fig. 6, 7) emphasizes the hips x (displacement on the sagittal plane, fig.2, 3). Principal Component 2 (PC2) highlights the combination 'right and left knee and shoulders' towards all directions (x, y, z). In fig. 6, 7 each box depicts the amount of energy over time, e.g. a joint is white only in the first half of the box, if it is highlighted only for the first half of the action.

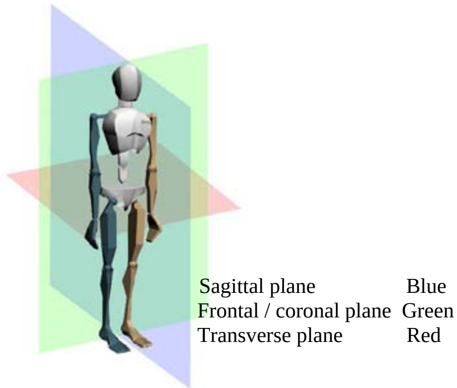

fig.3 X Moven Suit's body planes

The visualization presented here (fig.4, 5) supports the results of brain imaging studies; the schematic of the distributed semantic representation in the

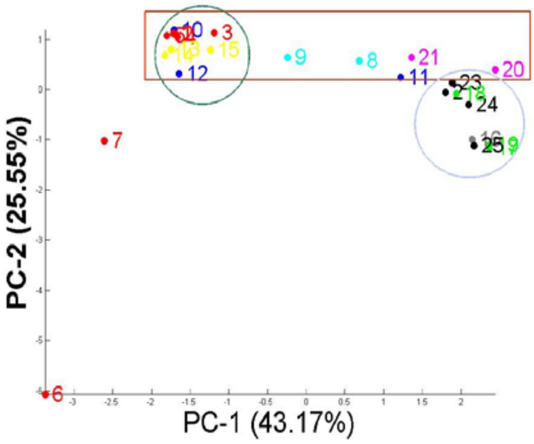

fig.4 Average Actions -Greek Verbs
(for numbers please seeTable 1)

brain of action verb processing is based on the body parts performing them (Wermter *et al.*, 2005), e.g. arm-related and leg-related.

Comparing this visualization with the linguistic classification (Antonopoulou, 1987; Levin; 1993) we see that Antonopoulou's classes are more fine-grained than the two big categories in fig.3, where all actions consisting each of the two linguistic groups appear in the same scheme –rectangular or ellipsis. Especially, the members of the walk-group, *pido*-group and *kateveno*-group are coiled together (green ellipsis). The same occurs for *girizo2*-group, *katevazo* and *anevazo* group (blue ellipsis). Levin's classes are similar to the groups at fig.5.

The actions 20, 21 (*girizo1*) and 22, 23, 24, 25 (*girizo2*) share the same morphological representation *girizo*. These actions present themselves close to each other on the PC1 projection but are separated on the PC2 projection fig.4). The first group -the green ellipsis- is mainly about leg motion. The second group -blue ellipsis- is about object manipulation that necessarily involves arm displacement as is clearly indicated on PC2. In English all *roll* verbs belong to the same class.

Though the plots were based on similar or even identical actions, certain divergences occurred. For instance, the verb *draskelizo* has traditionally been translated as *stride*, since both share longer steps. The Greek subjects always needed a small obstacle, such as a ball or a hole on the ground, to perform longer steps. On the contrary, the English subjects clearly distinguished between *stride* and *step over* (although WordNet assigns this meaning to *stride* as well). Similarly, we would expect *vimatizo* to be closer to *pace* rather than *march*, but it should be noted that the majority of the English participants were unsure for the exact representation of *pace*. When we compare the signals of *vimatizo* and *march*, we see significant similarity in the manner of stepping, while English subjects also emphasize the movement of the arms.

Of particular interest in the Greek plot is the distance of *treho*-*run* and *mpousoulo*-*crawl* from the rest leg-related actions (especially for English, *march* and *crawl* show the same behavior). Although we would expect *run* to resemble *treho*, differences occurred due to two reasons: (i) the English subjects tented to use their arms less than the Greek ones, and (ii) each group performed *march* in a different way; this time, the English subjects used their arms more than the Greek ones (as opposed to the performance of *run*). The blue and green ellipsis are projected both on the same PC1 (highlighted hips) and 2 (highlighted knees and shoulders).

Furthermore, the reason why *march*, *treho* and *crawl*-*mpousoulo* are projected on different PC2 narrows down to (i) the height of the knees (upward-downward and forward-backward respectively) and, (ii) the frequent movement of shoulders in all directions. But still, these verbs are projected on the same PC1. PC1 is about the forward and backward displacement of the hips and emphasizes on the leg related actions, namely the walking like actions, which, in turn, is considered to be the actions' common linguistic characteristic.

The above findings are still consistent with the aforementioned linguistic analysis. The Greek *treho* can form its own class in terms of velocity according to (Antonopoulou, 1983). At the same time, in fig.4, action 7 differs from all the other actions performed because of velocity. This fact is incorporated in the depiction of time in fig. 6. Probably Levin (1993) gives us a hint that these verbs need special treatment, since she enlists them under both the *meander* verbs and the *run* class (it must be kept in mind that

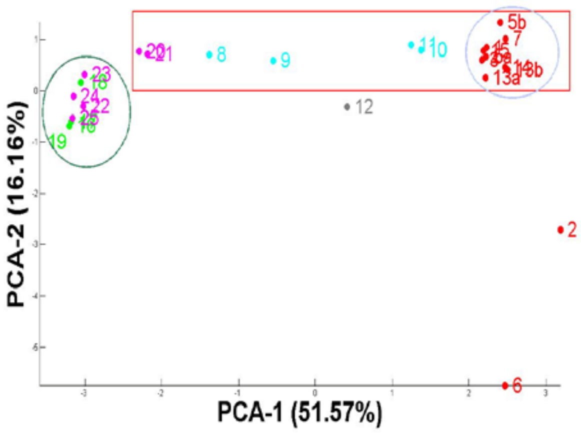

fig.5 Average Actions - English Verbs (colors according to Levin's classes)

probably Levin's classification has taken into account the criterion of intentionality, among others; however, intentionality is still not measurable with sensorimotor techniques).

Based on these first observations, we focus on our ongoing work that aims at proposing a framework that would establish joint-angle-based representations for parameters widely used in linguistic descriptions/classifications, such as path and directionality. The existing literature is mainly constricted in revealing the path of the action based on the gaze (Clark *et al.*, 2000). Intuitively, we could assume that directionality is mainly shown by the gross motion, e.g. head and chest movement or the upward and downward motion derived from the y axis of the knees, but further analysis needs to be done. Last but not least, we will extend our work by projecting the actions of each linguistic class separately and reduce their feature space, in order to focus on the most important synergies.

Table1 Verbs according to classes

Greek (Antonopoulou, 1987)			English (Levin, 1993)			English (FrameNet)	
Perpato verbs (treho can be the head of its own class)	Perpato	1	Run verbs (+meander)	Walk	1	Self motion	Walk
	Vimatizo	2		March	2		March
	Pisopato	3		Step back	3		Step back
	Triklizo	4		Stagger	4		Stagger
	Draskelizo	5		Stride/step over	5a/5b		Stride/step over
	Mpousoulo	6		Crawl	6		Crawl
	Treho	7		Run	7		Run
Aneveno verbs (upward motion)	Aneveno	8	Go verbs	Go up (step)	8	Motion	Go up (step)
	Aniforizo	9		Go up (ramp)	9		Go up (ramp)
Kateveno verbs (downward motion)	Kateveno	10		Go down (step)	10		Go down (step)
	Katiforizo	11		Go down (ramp)	11		Go down (ramp)
	Hamilono (only with the body)	12	Verb of assuming a position	Crouch	12	Posture	Crouch
Pido verbs	Pido (epi topou)	13	Run verbs	Jump/hop	13a/13b	Self motion	Jump/hop
	Pido (pano apo)	14		Jump over	14		Jump over
	Pido (apo kapou)	15		Jump down	15		Jump down
Katevazo verbs (downward motion)	Katevazo	16	Only the combination of the two verbs expresses the same with the Greek	Pick up and put on (lower)/ lower onto	16		Pick up and put on (lower)/ lower onto
Anevazo verbs (upward motion). Though sikono can form its own class	Anevazo	17		Pick up and put on (higher)/ lift onto	17		Pick up and put on (higher)/ lift onto
	Sikono	18	Lift verbs	Lift/raise	18	Body movement (raise not included in the same group)	Lift/raise
	Ipsono	19		Lift high	19		Lift high
Girizo$_1$verbs (rotary motion)	Girizo (antithetic katefthinsi)	20	Roll verbs (around an axis turn, rotate, circle)	Turn around	20	Change direction	Turn (as verb of changing direction)
	Girizo (e.g.giro apo karekla)	21		Circle (e.g. chair)	21	Motion	Circle (e.g. chair)
Girizo$_2$verbs (cause to turn)	Girizo (e.g. selida)	22		Turn (e.g. page)	22	Cause to move in place	Turn (as verb that cause to move in place)
	Peristrefo	23		Rotate	23		Rotate
	Anapodogirizo	24		Turn over	24		Turn over
	Kilo	25		Roll	25	Cause motion	Roll

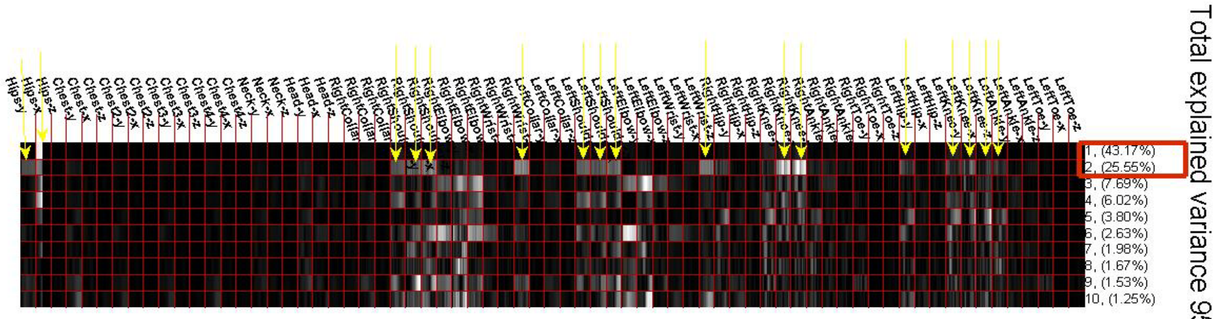

fig.6 Weights of each joint-direction-time feature according to the 10-first principal components (Greek Verbs)

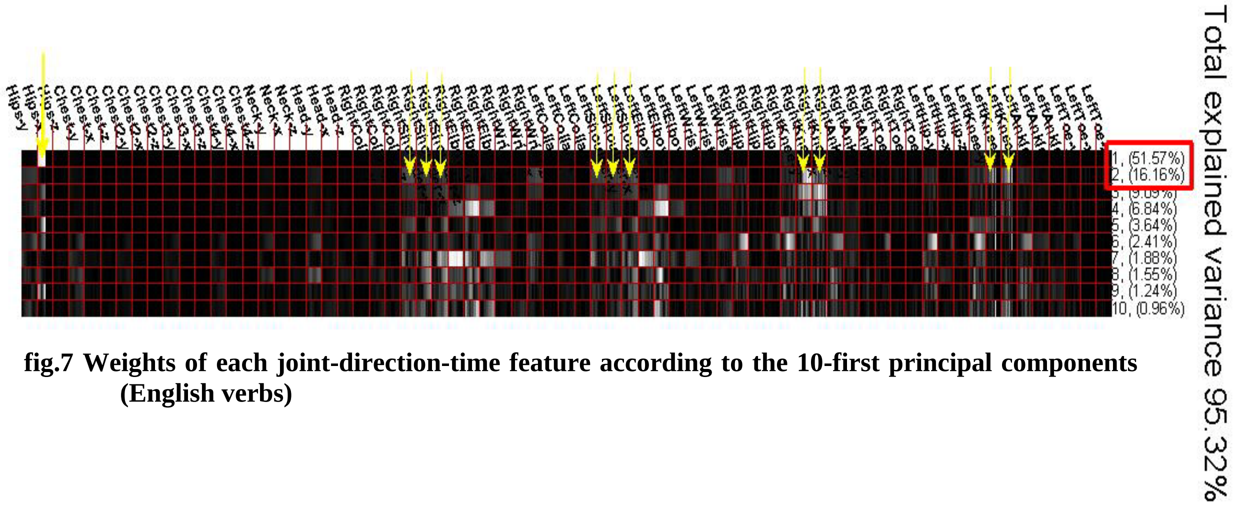

fig.7 Weights of each joint-direction-time feature according to the 10-first principal components (English verbs)

References

Antonopoulou, E. (1987). Prototype theory and the meaning of verbs, with special reference to Modern Greek verbs of motion. Unpublished PhD thesis. University of London

Arbib, M. A. (2008) From Grasp to Language: Embodied Concepts and the Challenge of Abstraction. Journal of Physiology Paris 102: pp. 4-20

Baker, C.F & Ruppenhofer, J. FrameNet's frames, vs, Levin's verbs classes.20-09-2010 <framenet.icsi.berkeley.edu/papers/bakerrup.pdf >

Clark, R.A., Miller, J.M. & Demer, J.L. (2000). Three-dimensional location of human rectus pulleys by path inflections in secondary gaze positions. Invest. Ophthalmol. &Visual Sci. 41

Fadiga, L., et al. (2006). Language in shadow. Social Neuroscience, 1(2): pp.77-89.

Feldman, J.(2006). From Molecule to Metaphor. Cambridge, MA: MIT Press.

Framenet http://framenet.icsi.berkeley.edu/

Galen, A, Grenager. T. & Manning C. (2004). Verb Sense and Subcategorization: Using Joint Inference to Improve Performance on Complementary Tasks. EMNLP 2004: pp. 150-157.

Jackendoff, R. (1990). Semantics Structures. Cambridge, MA: MIT Press.

Kemmerer, D. (2006). Action verbs, argument structure constructions, and the mirror neuron system. In M. Arbib (Ed.), Action to language via the mirror neuron system. Cambridge University Press.

Lakoff, G.& Johnson, M. (1999). Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought. Basic Books, New York

Levin, B. (1993). English Verb Classes and Alternations: A Preliminary Investigation, University of Chicago Press, Chicago, IL.

Papafragou, A., Massey, C., & Gleitman, L. (2006). When English proposes what Greek presupposes: The cross-linguistic encoding of motion events. Cognition 98: B75-87

Santello, M., Flanders, M., & Soechting, J.F. (1998). Postural hand synergies for tool use. Journal of Neuroscience, 18.

Troje, N.F. (2002). Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. Journal of Vision, 2: pp. 371–387

Wermter, S., Weber, C., Elshaw, M., Gallese, V. & Pulvermüller, F. (2005). Grounding neural robot language in action. Biomimetic Neural Learning for Intelligent Robots. Springer