# On Mental Imagery in Lexical Processing: Computational Modeling of the Visual Load Associated to Concepts

**Daniele P. Radicioni$^{\chi}$, Francesca Garbarini$^{\psi}$, Fabrizio Calzavarini$^{\phi}$**
**Monica Biggio$^{\psi}$, Antonio Lieto$^{\chi}$, Katiuscia Sacco$^{\psi}$, Diego Marconi$^{\phi}$**

(FirstName.Surname@unito.it)

$^{\chi}$Department of Computer Science, Turin University – Turin, Italy
$^{\phi}$Department of Philosophy, Turin University – Turin, Italy
$^{\psi}$Department of Psychology, Turin University – Turin, Italy

## Abstract

This paper investigates the notion of *visual load*, an estimate for a lexical item's efficacy in activating mental images associated with the concept it refers to. We elaborate on the centrality of this notion which is deeply and variously connected to lexical processing. A computational model of the visual load is introduced that builds on few low level features and on the dependency structure of sentences. The system implementing the proposed model has been experimentally assessed and shown to reasonably approximate human response.

**Keywords:** Visual imagery; Computational modeling; Natural Language Processing.

## Introduction

Ordinary experience suggests that lexical competence, i.e. the ability to use words, includes both the ability to relate words to the external world as accessed through perception (*referential* tasks) and the ability to relate words to other words in *inferential* tasks of several kinds (Marconi, 1997). There is evidence from both traditional neuropsychology and more recent neuroimaging research that the two aspects of lexical competence may be implemented by partly different brain processes. However, some very recent experiments appear to show that typically visual areas are also engaged by purely inferential tasks, not involving visual perception of objects or pictures (Marconi et al., 2013). The present work can be considered as a preliminary investigation aimed at verifying this main hypothesis, by investigating the following issues: *i)* to what extent the visual load associated with concepts can be assessed, and which sort of agreement exists among humans about the visual load associated to concepts; *ii)* which features underlie the visual load associated to concepts; and *iii)* whether the notion of visual load can be grasped and encapsulated into a computational model.

As it is widely acknowledged, one main visual correlate of language is *imageability*, that is the property of a particular word or sentence to produce an experience of imagery: in the following, we focus on *visual* imagery (thus disregarding *acoustic*, *olfactory* and *tactile* imagery), which we denote as *visual load*. The visual load is related to the easiness of producing visual imagery when an external linguistic stimulus is processed. Intuitively, words like 'dog' or 'apple' refer to *concrete* entities and are associated with a high visual load, implying that these terms immediately generate a mental image. Conversely, words like 'algebra' or 'idempotence' are hardly accompanied by the production of vivid images. Although the construct of visual load is closely related to that of concreteness, concreteness and visual load can clearly dissociate, in that *i)* some words have been rated high in visual load but low in concreteness, such as some concrete nouns that have been rated low in visual load (Paivio, Yuille, & Madigan, 1968); and, conversely, *ii)* abstract words such as 'bisection' are associated with a high visual load.

The notion of visual load is relevant to many disciplines, in that it contributes to shed light on a wide variety of cognitive and linguistic tasks and helps explaining a plethora of phenomena observed in both impaired and normal subjects. In the next Section we survey a multidisciplinary literature showing how mental imagery affects memory, learning and comprehension; we consider how imagery is characterized at the neural level; and we show how visual information is exploited in state-of-the-art Natural Language Processing research. In the subsequent Section we illustrate the proposed computational model for providing concepts with their visual load characterization. We then describe the experiments designed to assess the model through an implemented system, report and discuss the obtained results. Conclusion will summarize the work done and provide an outlook on future work.

## Related Work

As regards linguistic competence, it is generally accepted that visual load facilitates cognitive performance (Bergen, Lindsay, Matlock, & Narayanan, 2007), leading to faster lexical decisions than not-visually loaded concepts (Cortese & Khanna, 2007). For example, nouns with high visual load ratings are remembered better than those with low visual load ratings in long-term memory tests (Paivio et al., 1968). Moreover, visually loaded terms are easier to recognize for subjects with deep dyslexia, and individuals respond

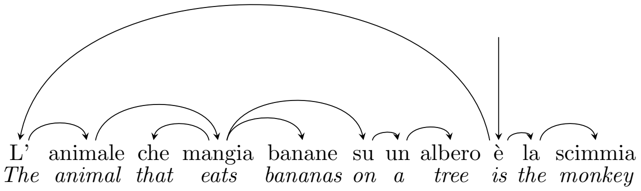
L' animale che mangia banane su un albero è la scimmia
The animal that eats bananas on a tree is the monkey

Figure 1: The (simplified) dependency tree corresponding to the sentence 'The animal that eats bananas on a tree is the Monkey'.

more quickly and accurately when making judgments about visually loaded sentences (Kiran & Tuchtenhagen, 2005). Neuropsychological research has shown that many aphasic patients perform better with linguistic items that more easily elicit visual imagery (Coltheart, 1980), although the opposite pattern has also been documented (Cipolotti & Warrington, 1995).

Visual imageability of *concepts* evoked by words and sentences is commonly known to affect brain activity. While visuosemantic processing regions, such as left inferior temporal gyrus and fusiform gyrus revealed greater involvement during the comprehension of highly imageable words and sentences (Bookheimer et al., 1998; Mellet, Tzourio, Denis, & Mazoyer, 1998), other semantic brain regions (i.e., superior and middle temporal cortex) are selectively activated by low-imageable sentences (Mellet et al., 1998; Just, Newman, Keller, McEleney, & Carpenter, 2004). Furthermore, a growing number of studies suggests that words encoding different visual properties (such as color, shape, motion, *etc.*) are processed in cortical areas that overlap with some of the areas that are activated during visual perception of those properties (Kemmerer, 2010).

Investigating the visual features associated to linguistic input can be useful to build semantic resources designed to deal with Natural Language Processing (NLP) problems, such as individuating verbs subcategorization frames (Bergsma & Goebel, 2011), enriching the traditional extraction of distributional semantics from text with a multimodal approach, integrating textual features with visual ones (Bruni, Tran, & Baroni, 2014). Finally, visual attributes are at the base of the development of annotated corpora and resources that can be used to extend text-based distributional semantics by grounding word meanings on visual features, as well (Silberer, Ferrari, & Lapata, 2013).

## Model

Although much work has been invested in different areas for investigating imageability in general and visual imagery in particular, at the best of our knowledge no attempt has been carried out to formally characterize visual load, and no computational model has been devised to compute how visually loaded are sentences and lexicalized concepts therein. We propose a model that relies on a simple hypothesis additively combining few low-level features, refined by exploiting syntactic information.

The notion of visual load, in fact, is used by and large in literature with different meanings, thus giving rise to different levels of ambiguity. We define visual load as the concept representing a direct indicator (a numeric value) of the efficacy for a lexical item to activate mental images associated to the concept referred to by the lexical item. We expect that visual load also represents an indirect measure of the probability of activation of brain areas deputed to the visual processing.

We conjecture that the visual load is primarily associated to *concepts*, although lexical phenomena like *terms availability* (implying that the most frequently used terms are easier to recognize than those seen less often (Tversky & Kahneman, 1973)) can also affect it.

Based on the work by Kemmerer (2010) we explore the hypothesis that a limited number of primitive elements can be used to characterize and evaluate the visual load associated to concepts. Namely, Kemmerer's Simulation Framework allows to grasp information about a wide variety of concepts and properties used to denote objects, events and spatial relations. Three main visual semantic components have been individuated that, in our opinion, are also suitable to be used as different dimensions along which to characterize the concept of visual load. They are: *color* properties, *shape* properties, and *motion* properties. The perception of these properties is expected to occur in a immediate way, such that "during our ordinary observation of the world, these three attributes of objects are tightly bound together in unified conscious images" (Kemmerer, 2010). We added a further perceptual component related to *size*. More precisely, our assumption is that information about the size of a given concept can also contribute, as an adjoint factor and not as a primitive one, to the computation of a visual load value for the considered concept.

In this setting, we represent each concept/property as a *boolean*-valued vector of four elements, each encoding the following information: *lemma*, morphological information on POS (part of speech), and then whether the considered concept/property conveys information about *color*, *shape*, *motion* and *size*.[1] For example, this piece of information

$$\mathsf{table,Noun,1,1,0,1} \tag{1}$$

can be used to indicate that the concept table (associated with a Noun, and differing, e.g., from that associated with a Verb) conveys information about color, shape and size, but not about motion. In the following, these are

[1] We adopt here a simplification, since we are assuming that the pair $\langle \text{lemma}, \text{POS} \rangle$ is sufficient to identify a concept/property, and that in general we can access items by disregarding the word sense disambiguation problem, which is known as an open problem in the field of NLP.

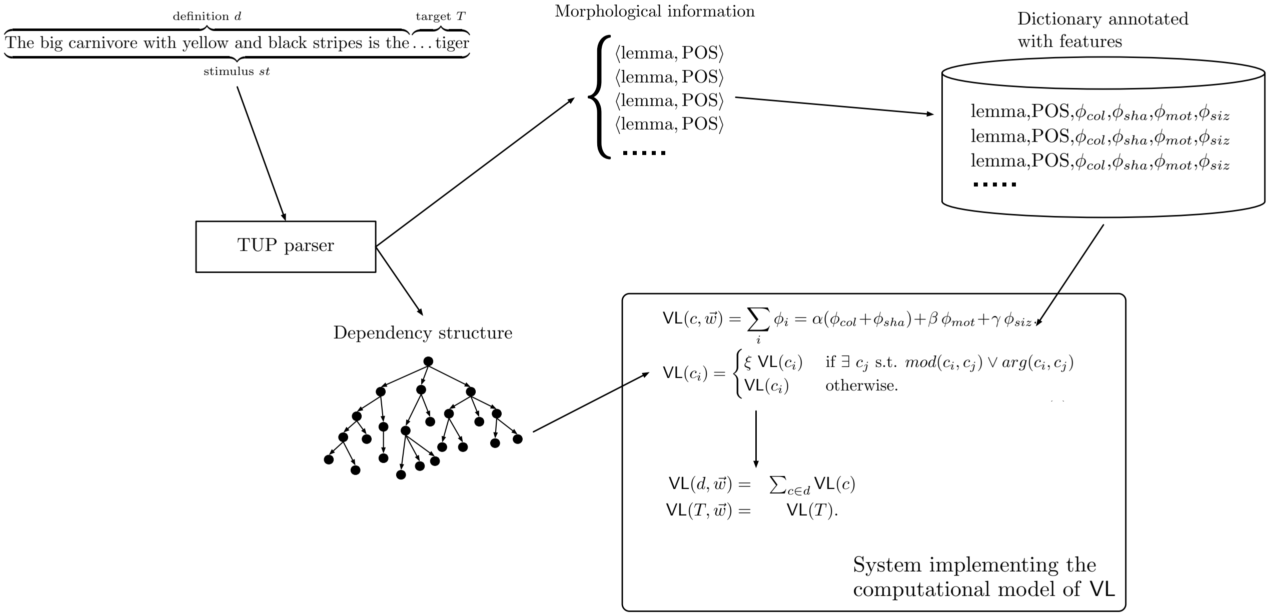
definition $d$
target $T$
The big carnivore with yellow and black stripes is the... tiger
stimulus $st$
TUP parser
Morphological information
$\langle$lemma, POS$\rangle$
$\langle$lemma, POS$\rangle$
$\langle$lemma, POS$\rangle$
$\langle$lemma, POS$\rangle$
.....
Dictionary annotated with features
lemma,POS,$\phi_{col}$,$\phi_{sha}$,$\phi_{mot}$,$\phi_{siz}$
lemma,POS,$\phi_{col}$,$\phi_{sha}$,$\phi_{mot}$,$\phi_{siz}$
lemma,POS,$\phi_{col}$,$\phi_{sha}$,$\phi_{mot}$,$\phi_{siz}$
.....
Dependency structure
$\mathsf{VL}(c,\vec{w}) = \sum_i \phi_i = \alpha(\phi_{col}+\phi_{sha})+\beta\,\phi_{mot}+\gamma\,\phi_{siz}$
$\mathsf{VL}(c_i) = \begin{cases}\xi\ \mathsf{VL}(c_i) & \text{if } \exists\ c_j \text{ s.t. } mod(c_i,c_j) \vee arg(c_i,c_j)\\ \mathsf{VL}(c_i) & \text{otherwise.}\end{cases}$
$\mathsf{VL}(d,\vec{w}) = \sum_{c\in d}\mathsf{VL}(c)$
$\mathsf{VL}(T,\vec{w}) = \mathsf{VL}(T).$
System implementing the computational model of VL

Figure 2: The pipeline to compute the $\mathsf{VL}$ score according to the proposed computational model.

referred to as the visual features $\phi$. associated with the given concept.

We have then built a dictionary by extracting it from a set of *stimuli* (illustrated hereafter) composed of simple sentences describing a concept; next, we have manually annotated the visual features associated with each concept. The automatic annotation of visual properties associated with concepts is deferred to future work: it can be addressed either through a classical Information Extraction approach building on statistics, or in a more semantically-principled way.

Different weighting schemes $\vec{w} = \{\alpha, \beta, \gamma\}$ have been tested in order to determine the features' contribution to the visual load associated with a concept $c$, that results from computing

$$\mathsf{VL}(c,\vec{w}) = \sum_i \phi_i = \alpha(\phi_{col}+\phi_{sha})+\beta\,\phi_{mot}+\gamma\,\phi_{siz}. \quad (2)$$

For the experimentation we set $\alpha$ to 1.35, $\beta$ to 1.1 and $\gamma$ to .9: these assignments reflect the fact that color and shape information is considered more important, in the computation of $\mathsf{VL}$.

To the ends of combining the contribution of concepts in a sentence $s$ to the overall $\mathsf{VL}$ score for $s$, we adopted the following additive schema: $\mathsf{VL}(s) = \sum_{c\in s}\mathsf{VL}(c)$.

The computation of the $\mathsf{VL}$ score also accounts for the dependency structure of the input sentences. The syntactic structure of sentences is computed by the Turin University Parser (TUP) in the dependency format (Lesmo, 2007). Dependency formalisms represent syntactic relations by connecting a *dominant* word, the head (e.g., the verb 'fly' in the sentence *The eagle flies*) and a *dominated* word, the dependent (e.g., the noun 'eagle' in the same sentence). The connection between these two words is usually represented by using labeled directed edges (e.g., *subject*): the collection of all dependency relations of a sentence forms a tree, rooted in the main verb (see the parse tree illustrated in Figure 1). The dependency structure is relevant in our approach, because we assume that a *reinforcement* effect may apply in cases where both a word and its dependent(s) (or governor(s)) are associated with visual features. For example, a phrase such as 'with black stripes' is expected to evoke mental images in a more vivid way than its elements taken in isolation (that is, 'black' and 'stripes'), moreover its visual load is expected to further grow if we add a coordinated term, as in 'with *yellow and* black stripes'. Moreover, the $\mathsf{VL}$ would –recursively– grow if we added a governor term (like '*fur* with yellow and black stripes'). We then introduced a parameter $\xi$ to control the contribution of the aforementioned features in case the corresponding terms are linked in the parse tree by a modifier/argument relation (denoted as *mod* and *arg* in Equation 3).

$$\mathsf{VL}(c_i) = \begin{cases}\xi\ \mathsf{VL}(c_i) & \text{if } \exists\ c_j \text{ s.t. } mod(c_i,c_j) \vee arg(c_i,c_j)\\ \mathsf{VL}(c_i) & \text{otherwise.}\end{cases} \quad (3)$$

In the experimentation $\xi$ was set to 1.2.

The stimuli in the dataset are pairs consisting of a definition $d$ and a target $T$ ($st = \langle d, T\rangle$), such as

$$\underbrace{\overbrace{\text{The big carnivore with yellow and black stripes is the}}^{\text{definition } d}\overbrace{\ldots\text{tiger}}^{\text{target } T}}_{\text{stimulus } st}.$$

The visual load associated to $st$ components, given the weighting scheme $\vec{w}$, is then computed as follows:

$$\mathsf{VL}(d, \vec{w}) = \sum_{c \in d} \mathsf{VL}(c) \tag{4}$$

$$\mathsf{VL}(T, \vec{w}) = \mathsf{VL}(T). \tag{5}$$

The whole pipeline from the input parsing to computation of the VL for the considered stimulus has been implemented as a computer program; its main steps include the parsing of the stimulus, the extraction of the (lexicalized) concepts by exploiting the output of the morphological analysis, and the tree traversal of the dependency structure resulting from the parsing step. The morphological analyzer has been preliminarily fed with the whole set of stimuli, and its output has been annotated with the visual features and stored into a dictionary. At run time, the dictionary is accessed based on morphological information, then used to retrieve the values of the features associated with the concepts in the stimulus. The output obtained by the proposed model has been compared with the results obtained in a behavioral experimentation as described below.

# Experimentation

## Materials and Methods

Thirty healthy volunteers, native Italian speakers, (16 females and 14 males), $19 - 52$ years of age (mean $\pm sd = 25.7 \pm 5.1$), were recruited for the experiment. None of the subjects had a history of psychiatric or neurological disorders. All participants gave their written informed consent before participating in the experimental procedure, which was approved by the ethical committee of the University of Turin, in accordance with the Declaration of Helsinki (World Medical Association, 1991). Participants were all naïve to the experimental procedure and to the aims of the study.

**Experimental design and procedure** Participants were asked to perform an inferential task "Naming from definition". During the task a sentence was pronounced and the subjects were instructed to listen to the stimulus given in the headphones and to overtly name, as accurately and as fast as possible, the target word corresponding to the definition, using a microphone connected to a response box. Auditory stimuli were presented through the E-Prime software, which was also used to record data on accuracy and reaction times. Furthermore, at the end of the experimental session, the subjects were administered a questionnaire: they had to rate on a $1 - 7$ Likert scale the intensity of the visual load they perceived as related to each target and to each definition.

The factorial design of the study included two within-subjects factors, in which the visual load of both target and definition was manipulated. The resulting four experimental conditions were as follows:

**VV** Visual Target—Visual Definition (e.g., 'The bird of prey with great wings flying over the mountains is the ... eagle');

**VNV** Visual Target—Non-Visual Definition (e.g., The hottest of the four elements of the ancients is ... fire);

**NVV** Non-Visual Target—Visual Definition (e.g., The nose of Pinocchio stretched when he told a ... lie);

**NVNV** Non-Visual Target—Non-Visual Definition (e.g., The quality of people that easily solve difficult problems is said ... intelligence).

For each condition, there were 48 sentences, 192 sentences overall. Each trial lasted about 30 minutes. The number of words (nouns and adjectives), their balancing across stimuli, and the (syntactic dependency) structure of the considered sentences were uniform within conditions, so that the most relevant variables were controlled. The same set of stimuli used for the human experiment was given in input to the system implementing the computational model.

## Data analysis

The participants' performance in the "Naming from definition" task was evaluated by recording, for each response, the reaction time *RT*, in milliseconds, and the accuracy *AC*, computed as the percentage of correct answers. The answers were considered *correct* if the target word was plausibly matched with the definition. Then, for each subject, both *RT* and *AC* were combined in the *Inverse Efficiency Score* (IES), by using the formula IES $= (RT/AC) \cdot 100$. IES is a metrics commonly used to aggregate reaction time and accuracy, and to summarize them (Townsend & Ashby, 1978). The mean IES value was used as the dependent variable and entered in a $2 \times 2$ repeated measures ANOVA with 'target' (two levels: 'visual' and 'non-visual') and 'definition' (two levels: 'visual' and 'non-visual') as within-subjects factors. *Post hoc* comparisons were performed by using the Duncan test.

The scores obtained by the participants in the visual load questionnaire were analyzed by using unpaired T-tests, two tailed. Two comparisons were performed for visual and non-visual targets, and for visual and non-visual definitions. The computational model results were analyzed by using unpaired T-tests, two tailed. Two comparisons were performed for visual and non-visual targets and for visual and non-visual definitions.

**Correlations between IES, computational model and visual load questionnaire**. We also explored the existence of correlations between IES, the visual load questionnaire and the computational model output by using linear regressions. For both the IES values and the questionnaire scores, we computed for each item the mean of the 30 subjects' responses. In a first model, we used the visual load questionnaire scores as independent variable to predict the participants' performance (with

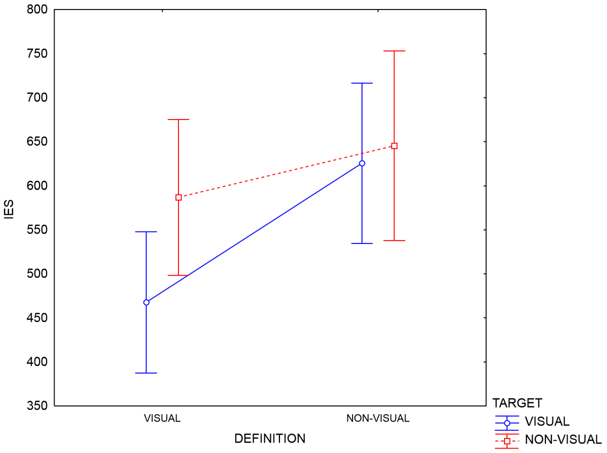

Figure 3: The graph shows, for each condition, the mean IES with standard error.

IESas the dependent variable); in a second model, we used the computational data as independent variable to predict the participants' visual load evaluation (with the questionnaire scores as the independent variable). In order to verify the consistency of the correlation effects, we also performed linear regressions where we controlled for three covariate variables: the number of words, their balancing across stimuli and the syntactic dependency structure.

## Results

The ANOVA showed a significant effect of the within-subject factors "target" ($F_{1,29} = 14.4$; $p < 0.001$), suggesting that the IES values were significantly lower in the visual than in the non-visual targets, and "definition" ($F_{1,29} = 32.78$; $p < 0.001$), suggesting that the IES values were significantly lower in the visual than in the non-visual definitions. This means that, for both the target and the definition, the participants' performance was significantly faster and more accurate in the visual than in the non-visual condition. We also found a significant interaction "target*definition" ($F_{1,29} = 7.54$; $p = 0.01$). Based on the Duncan *post hoc* comparison, we verified that this interaction was explained by the effect of the visual definitions of the visual targets (VV condition), in which the participants' performance was significantly faster and more accurate than in all the other conditions (VNV; NVV; NVNV), as shown in Figure 3.

By comparing the questionnaire scores for visual (mean $\pm sd = 5.69 \pm 0.55$) and non-visual (mean $\pm sd = 4.73 \pm 0.71$) definitions we found a significant difference ($p < 0.001$; unpaired T-test, two tailed). By comparing the questionnaire scores for visual (mean $\pm sd = 6.32 \pm 0.4$) and non-visual (mean $\pm sd = 4.23 \pm 0.9$) targets we found a significant difference ($p < 0.001$). This suggest that our arbitrary categorization of each sentences within the four conditions was supported by the general agreement of the subjects. By comparing the computational model scores for visual (mean $\pm sd = 4.0 \pm 2.4$) and non-visual (mean $\pm sd = 2.9 \pm 2.0$) definitions we found a significant difference ($p < 0.001$; unpaired T-test, two tailed). By comparing the computational model scores for visual (mean $\pm sd = 2.53 \pm 1.29$) and non-visual (mean $\pm sd = 0.26 \pm 0.64$) targets we found a significant difference ($p < 0.001$). This suggest that we were able to computationally model the visual-load of both targets and descriptions, describing it as a linear combination of different low-level features: color, shape, motion and dimension.

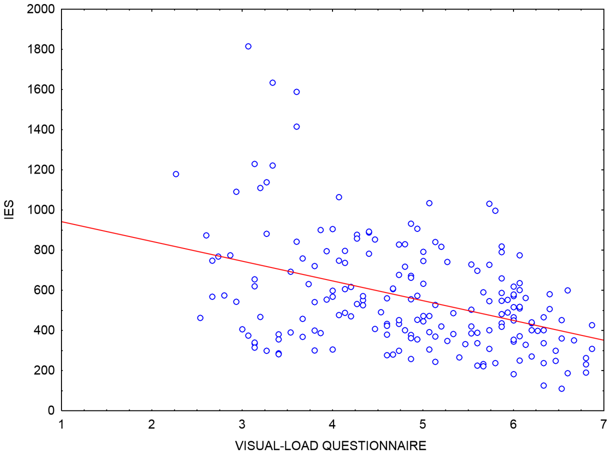

Figure 4: Linear regression "Inverse Efficiency Score (IES) by Visual Load Questionnaire". The mean score in the Visual Load Questionnaire, reported on $1-7$ Likert scale, was used as an independent variable to predict the subjects' performance, as quantified by the IES.

**Results correlations**. By using the visual load questionnaire scores as independent variable we were able to significantly ($R^2 = 0.4; p < 0.001$) predict the participants' performance (that is, their IES values), illustrated in Figure 4. This means that the higher the participants' visual score for a definition, the better the participants' performance in giving the correct response (or, alternatively, the lower the IES value).

By using the computational data as independent variable we were able to significantly ($R^2 = 0.44; p < 0.001$) predict the participants' visual load evaluation (their questionnaire scores), as shown in Figure 5. This means that a correlation exists between the computational prediction about the visual load of the definitions and the participants visual load evaluation: the higher is the computational model result, the higher is the participants' visual score in the questionnaire. We also found that these effects were still significant in the regression models where the number of words, their balancing across stimuli and the syntactic dependency structure was controlled for.

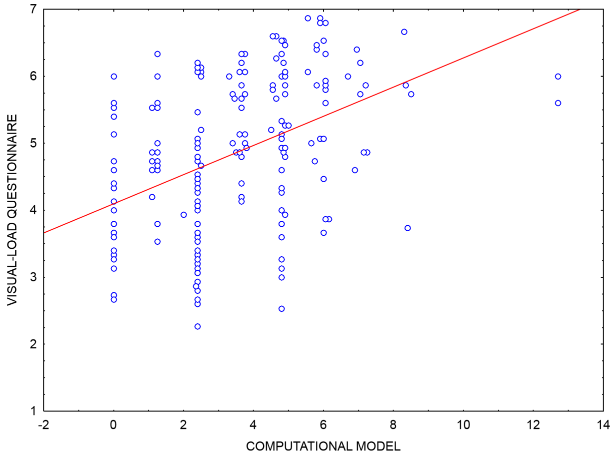

Figure 5: Linear regression "Visual Load Questionnaire by Computational Model". The mean value obtained by the Computational model was used as an independent variable to predict the subjects' scores on the Visual Load Questionnaire, reported on $1 - 7$ Likert scale.

## Conclusions

In the next future we plan to extend the representation of the conceptual information by grounding the conceptual representation on a hybrid representation composed of conceptual spaces and ontologies (Lieto, Minieri, Piana, & Radicioni, 2015; Lieto, Radicioni, & Rho, 2015). Additionally, we plan to integrate the current model in the context of cognitive architectures.

## Acknowledgments

This work has been partly supported by the Project *The Role of the Visual Imagery in Lexical Processing*, grant TO-call03-2012-0046, funded by Università degli Studi di Torino and Compagnia di San Paolo.

## References

Bergen, B. K., Lindsay, S., Matlock, T., & Narayanan, S. (2007). Spatial and linguistic aspects of visual imagery in sentence comprehension. *Cognitive Sci*, *31*(5), 733–764.

Bergsma, S., & Goebel, R. (2011). Using visual information to predict lexical preference. In *RANLP* (pp. 399–405).

Bookheimer, S., Zeffiro, T., Blaxton, T., Gaillard, W., Malow, B., & Theodore, W. (1998). Regional cerebral blood flow during auditory responsive naming: evidence for cross-modality neural activation. *Neuroreport*, *9*(10), 2409–2413.

Bruni, E., Tran, N.-K., & Baroni, M. (2014). Multimodal distributional semantics. *J. Artif. Intell. Res.*, *49*, 1–47.

Cipolotti, L., & Warrington, E. K. (1995). Semantic memory and reading abilities: A case report. *J INT NEUROPSYCH SOC*, *1*(01), 104–110.

Coltheart, M. (1980). Deep dyslexia: A right hemisphere hypothesis. *Deep dyslexia*, 326–380.

Cortese, M. J., & Khanna, M. M. (2007). Age of acquisition predicts naming and lexical-decision performance above and beyond 22 other predictor variables: An analysis of 2,342 words. *Q J Exp Psychol A*, *60*(8), 1072–1082.

Just, M. A., Newman, S. D., Keller, T. A., McEleney, A., & Carpenter, P. A. (2004). Imagery in sentence comprehension: an fmri study. *Neuroimage*, *21*(1), 112–124.

Kemmerer, D. (2010). Words and the Mind: How words capture human experience. In B. Malt & P. Wolff (Eds.), (chap. How Words Capture Visual Experience - The Perspective from Cognitive Neuroscience). Oxford Scholarship Online.

Kiran, S., & Tuchtenhagen, J. (2005). Imageability effects in normal spanish–english bilingual adults and in aphasia: Evidence from naming to definition and semantic priming tasks. *Aphasiology*, *19*(3-5), 315–327.

Lesmo, L. (2007, June). The Rule-Based Parser of the NLP Group of the University of Torino. *Intelligenza Artificiale*, *2*(4), 46–47.

Lieto, A., Minieri, A., Piana, A., & Radicioni, D. P. (2015). A knowledge-based system for prototypical reasoning. *Connection Science*, *27*(2), 137–152.

Lieto, A., Radicioni, D. P., & Rho, V. (2015, July). A Common-Sense Conceptual Categorization System Integrating Heterogeneous Proxytypes and the Dual Process of Reasoning. In *Proc. of IJCAI 2015.* Buenos Aires, Argentina: AAAI Press.

Marconi, D. (1997). *Lexical competence.* MIT Press.

Marconi, D., Manenti, R., Catricala, E., Della Rosa, P. A., Siri, S., & Cappa, S. F. (2013). The neural substrates of inferential and referential semantic processing. *Cortex*, *49*(8), 2055–2066.

Mellet, E., Tzourio, N., Denis, M., & Mazoyer, B. (1998). Cortical anatomy of mental imagery of concrete nouns based on their dictionary definition. *Neuroreport*, *9*(5), 803–808.

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of experimental psychology*, *76*, 1.

Silberer, C., Ferrari, V., & Lapata, M. (2013). Models of semantic representation with visual attributes. In *Acl 2013 proceedings* (pp. 572–582).

Townsend, J. T., & Ashby, F. G. (1978). Methods of modeling capacity in simple processing systems. *Cognitive theory*, *3*, 200–239.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive psychology*, *5*(2), 207–232.

World Medical Association. (1991). Code of Ethics: Declaration of Helsinki. *BMJ*, *302*, 1194.