



Lorena Bort-Mir\*, Marianna Bolognesi and Susan Ghaffaryan

# Cross-cultural interpretation of filmic metaphors: A think-aloud experiment

<https://doi.org/10.1515/ip-2020-4001>

ISSN (printed): 1612-295X.

ISSN (electronic): 1613-365X.

**Abstract:** The purpose of this study is to investigate how viewers who speak different languages interpret cinematographic metaphors in a filmic advertisement. The study is organized in three parts: First, we offer a theoretical model that predicts the offline mental mechanisms that occur while people interpret filmic metaphors, based on an existing model of visual metaphor processing. Second, we evaluate the model in a think-aloud retrospective task. A TV-commercial is projected individually to 30 Spanish, 30 American, and 30 Persian participants, who are then asked to verbalize their thoughts. The commercial was previously segmented, analyzed using FILMIP (Filmic Metaphor Identification Procedure), and marked for metaphoricality by two independent analysts. The collected data is then evaluated in two formal content analyses. In the first one, two independent coders classified all the clauses used by the 90 participants in relation to the steps outlined in the theoretical model. In the second analysis, those clauses in which the participants were constructing their metaphorical interpretation of the filmic advertisement were annotated for the type of metaphor they constructed. The general results show that: (1) some mental processes seem to be more prominent in some cultures and not in others, and (2) genre-related knowledge plays a crucial role in constructing filmic metaphors in certain cultures and not in others. With this study, we theoretically formalize and empirically test the types of operations reflected in the language that viewers use to describe how they interpret filmic metaphors, thus advancing the current theory and methods on filmic metaphor interpretation from cognitive, semiotic, and cross-cultural perspectives.

**Keywords:** metaphor processing, think aloud method, filmic metaphor, metaphor construction, filmic narrative

---

\*Corresponding author: Lorena Bort-Mir, Universitat Jaume I, Castelló de la Plana, Spain, E-mail: [bortl@uji.es](mailto:bortl@uji.es)

Marianna Bolognesi, University of Bologna, Bologna, Italy, E-mail: [marianna.bolognesi@gmail.com](mailto:marianna.bolognesi@gmail.com)

Susan Ghaffaryan, Allameh Tabataba'i University, Tehran, Islamic Republic of Iran, E-mail: [susan.ghaffaryan@gmail.com](mailto:susan.ghaffaryan@gmail.com)

# 1 Introduction

Richards (1965: 94) proposed that “thought is metaphoric, and proceeds by comparison, and the metaphors of language derive therefrom.” If our thought is shaped by metaphors, then we are likely to find these tropes not only in language but also in other modes of expression. The study of not-just-linguistic metaphors has attracted high levels of interest in recent years. Many researchers (Forceville and Uriós-Aparisi 2009; Foss 2005; Phillips and McQuarrie 2004), for instance, address the way a metaphor is depicted in still images. Among the various types of pictorial genres, advertising has been widely investigated in relation to metaphorical constructions (e.g., Forceville 1996; McQuarrie and Mick 1999; Pérez-Sobrino 2017; Phillips 1997). The problem that typically remains neglected is how these pictorial metaphors can be identified in the wild among millions of still images, as well as how they are processed in the minds of the viewers. Šorm and Steen (2013) proposed a theoretical model of visual metaphor processing, with the identification of three broad categories involved in the processing of visual metaphors, which include (i) incongruity perception, (ii) incongruity resolution, and (iii) contextual processing.

A related means of communication that makes use of metaphors to express meaning is cinema. Coëgnarts and Kravanja (2012) conclude that abstract meaning is expressed in films through conceptual metaphors. According to them, “film-makers use embodied principles in the form of image schemas and conceptual metaphors to express abstract meaning to the spectator” (2012: 3). This construal of meaning by means of cinematic metaphors is also the focus of research carried out by Kappelhoff and Müller (2011) and Ortiz (2014), thus adding valuable insight into the study of how metaphors work in the filmic medium. However, there is still a need for more research to determine what type of operations viewers deploy to identify, construct and interpret metaphors in filmic materials.

The purpose of our study is to fill this gap by offering a theoretical model of filmic metaphor construction. Our model derives from two previous studies: (i) the model of visual metaphor processing proposed by Šorm and Steen (2013), and (ii) a cognitive processing model of aesthetic experiences of visual artefacts (Leder et al. 2004). The present model is then validated through a think-aloud retrospective task that constitutes our tool to collect verbal data on filmic metaphor construction. A TV commercial previously marked for metaphoricity with FILMIP<sup>1</sup> (Filmic Metaphor Identification Procedure, Bort-Mir 2019) was projected individually to 90 participants belonging to different cultures, and they were asked to verbalize their

---

**1** FILMIP is a structural method for the identification of filmic metaphors based on VISMIP, the Visual Metaphor Identification Procedure (Šorm and Steen 2018).

thoughts on it while being audio recorded. After the transcription process, all the data were segmented into clauses, which were later coded by independent annotators according to the distinct mental operations detected in the filmic metaphor construction. By means of content analyses and interrater reliability testing, we observed the different types of mental operations entertained by each group of participants (Spanish, American, and Persian participants), which we compared in a cross-cultural manner. Finally, we explored the type of metaphors identified and constructed by the different groups, in relation to the same filmic advertisement<sup>2</sup>.

Thus, our study addresses the following research questions:

1. Is it possible to determine a list of mental operations that reflect the way in which speakers of different languages describe, analyze and interpret metaphorical filmic advertisements?
2. Are there mental operations that are particularly related to specific groups of speakers?
3. Do speakers of different languages tend to identify and construct the same metaphors, when exposed to the same metaphorical filmic advertisement?
4. Can these offline cognitive processes be integrated into a theoretical model of filmic metaphor construction?

The results of the present paper are intended to empirically validate our model for offline<sup>3</sup> filmic metaphor construction.

## 2 Theoretical background

### 2.1 Visual metaphor processing and aesthetic processing

The model of offline filmic metaphor construction that we develop in this paper is based on the model of visual metaphor processing proposed by Šorm and Steen (2013), which was elaborated through a think-aloud study. In their experiment, 24 participants had to look at 20 images from different genres (political cartoons, advertisements, educational illustrations, and paintings) containing visual metaphors, and then they were asked to verbalize their thoughts. A subset of the

---

<sup>2</sup> With filmic advertisement we mean advertisements that are expressed by means of moving pictures. This is contrasted with print advertisements, a terminology typically used to identify advertisements expressed through still images (e.g., billboards).

<sup>3</sup> By offline processing, we refer to an indirect measurement of the way in which the speakers approach the processing of the filmic metaphor. Future experiments may focus on online processing strategies, using psycholinguistic techniques such as eye-tracking.

collected data was segmented into clauses and was annotated using a specific coding scheme that encompassed various categories. The researchers identified three broad categories of visual metaphor processing:

- (i) *Incongruity perception*: participants identify the components of the image and relate them to the fact that something seems strange to them. In this stage, participants analyze simple perceptual components such as colors, shapes, and objects. They recognize the scene they see, and then they realize that something is weird there, thus identifying the visual incongruity within the picture.
- (ii) *Incongruity resolution*: in this stage, participants try to resolve the identified problem (incongruity) in the previous phase by inferencing meaning from the mappings from the source and target domains. In this process, participants construct, interpret, and judge the metaphor.
- (iii) *Contextual processing*: participants express their thoughts about all the other information of the picture that can influence their interpretation (contextual details). They talk about the specific genre of the image, the author, or even about its historical context.

A theoretical model for the cognitive processing of metaphorical images was then developed based on an existing model of aesthetic appreciation (Leder et al. 2004).

Aesthetic experiences were thoroughly studied in 2004 under the eye of psycholinguistics by Leder, Belke, Oeberst, and Augustin in their *A model of aesthetic appreciation and aesthetic judgments*. The model describes the cognitive stages that people undertake while watching visual artistic stimuli, although it was also clearly stated that this model could also be transferred to other kinds of aesthetic experiences and visual stimuli. For our study, this assumption poses a question that must be solved before developing a model of offline filmic metaphor construction: is film an aesthetic experience, so that our model can also be based on Leder's model?

The genre of film is becoming very popular in distinct research fields and among the non-academic public, and scholars have recently postulated some theories about the aesthetic experience of this medium (Goldstein 2009; Grodal 2009; Hilscher et al. 2008). Marković (2012) investigates the components of aesthetic experience, distinguishing between appraisal, fascination, and emotion. According to the author, "every object of aesthetic processing has some physical form which determines the stylistic aspect of the artwork's identity. An aesthetic form is a specific composition of various features such as colors, lines, shapes, sounds, gestures, and so on" (Marković 2012: 9).

He also identified two analog stages of aesthetic information processing: (i) the narrative (the thematic and symbolic meaning of an artwork), and (ii) the composition (stylistic form of expressing an artwork) (Marković 2012: 10).

According to these assumptions, we assume that the filmic medium is considered an aesthetic experience, since films are complex gestalts of significance, composed of different layers of meaning expressed through different modes of communication (visuals, non-verbal sounds, spoken discourse, written discourse, and music, Forceville and Uriós-Aparisi 2009). Narrative in films is then expressed by means of a specific filmic composition (the *mise-en-scène*).

The processing of aesthetic experiences entails several distinct cognitive and affective processes (see Figure 1 below), all of them identified in the model proposed by Leder’s et al. (2004).

The first process, *perceptual analysis*, entails the perception of simple perceptual features of the artwork under analysis.

The second process is called *implicit memory integration*. This is an unconscious stage where participants match what they see with a set of conceptual representations stored in memory.

The third process, *explicit classification*, is affected by the knowledge or level of expertise of the viewers. This stage encompasses the analysis of both the content and the style of the artwork, thus responding to the questions “what is depicted?” and “how is it expressed?”.

*Cognitive mastering* is the fourth process identified by this model, and it entails the interpretation of the artwork by the viewers. Their level of expertise or their own feelings and personal experiences may also influence this interpretation.

Finally, during the *evaluation* process, the viewer’s judge the artwork, thus ending up in a process of aesthetic judgment and emotional processing, where they may have positive or negative feelings about what they see.

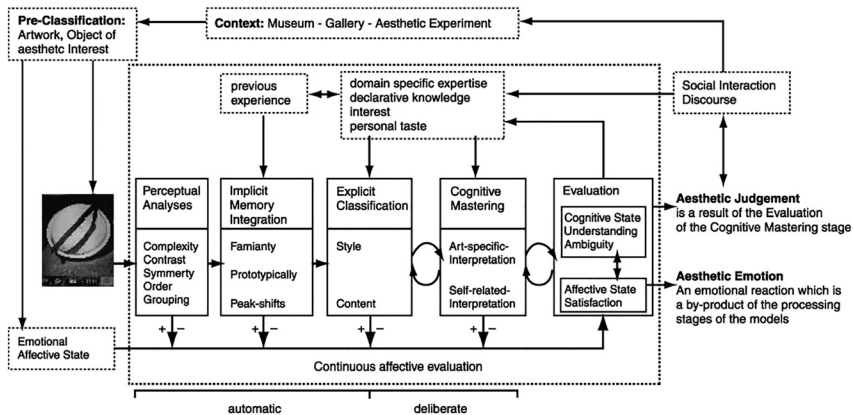


Figure 1: Leder’s et al. model of aesthetic experience (2004: 492).

This model of visual artefact processing (Leder et al. 2004) was integrated into the model of metaphor processing proposed by Šorm and Steen (2013). Our purpose now is to determine whether they can also be integrated into a model of the offline interpretation and narrative construction of metaphors in films.

## 2.2 A model of filmic metaphor construction

We propose an integration of Šorm and Steen's model (2013) and Leder's et al. model (2004) into our model of filmic metaphor construction. Such integration is made under the following premises:

First, we expect that participants go through a process of description of what they perceive from the screen, this is, what they see and hear within the clip. This process is encoded within the incongruity perception process in Šorm and Steen's model (2013). They claim that the description of simple perceptual elements goes along with the identification of incongruous elements within the picture. However, as the filmic medium is a highly complex means of communication where lots of distinct components are put together in order to convey meaning, the process of describing first what is perceived requires its inclusion into a separate category, which we call the *content description* process. This first process is connected to Leder's et al. (2004) stages of perceptual analyses (perception of simple perceptual elements) and explicit classification (description of content and style).

It is also within this first process of content description where we expect participants to identify and describe the particular genre of the film they are watching, an idea supported by Leder et al. (2004), who suggested that the contextual information of an artwork might indeed affect its understanding. This description of the genre, which puts at work the genre knowledge of the viewer and contextual information such as titles or knowledge of the brand or designer, would correspond to Leder's et al. pre-classification process and to Šorm and Steen's contextual processing stage.

Second, after this first descriptive phase, we expect a process where the metaphor interpretation and construction in itself takes place, thus involving not only the identification of incongruous elements within the film but also its resolution. Incongruity related to metaphor has been defined as the disparity between the contextual meaning of a discourse unit and its basic meaning (Steen 2007), thus involving a need for comparison. We assume that the identification of strange elements that do not fit well into the filmic narrative makes the viewer's try to find other components that allow for that comparison, which naturally leads to resolving that incongruity. This metaphor construction phase includes, then, Šorm and Steen's incongruity perception and incongruity resolution processes, and it

also encompasses metaphor recognition. Leder's et al. cognitive mastering stage would correspond to our metaphor construction category, as it entails the viewer's effort to understand the clip, making it meaningful in some way or another.

Finally, we propose that a stage where the clip is evaluated occurs after the metaphor is constructed. This process of evaluation of the visual stimulus is included within the incongruity resolution phase in Šorm and Steen's model (2013), but it is differentiated from the cognitive mastering in Leder's et al. model (2004), distinguishing between aesthetic judgment and emotional processing within the evaluation stage. We also assume this distinction for our model of offline filmic metaphor construction; however, the mechanisms of the genre under which our model is developed (TV advertising) demands a slight modification of this evaluative stage, consequently distinguishing among a process of communicative or message detection, and a process of appreciation of the clip.

We conclude that our model includes three broad categories for offline metaphor construction in films:

1. *Content description* involves the recognition of perceptual elements (colors, images, objects, people and their actions, text, music, spoken discourse, kinetics, etc.). The process of inferencing meaning starts here, which also involves the identification of the genre and all the contextual information that might be valuable to that meaning making of the film.
2. *Metaphor construction* includes the process of finding incongruous elements and resolving them with a comparison, thus leading to the recognition of the metaphor and its construction. We expect that not all the participants explicitly recognize a metaphor in the clip, as the labeling of the trope may be influenced by their level of expertise and knowledge of the world.
3. *Evaluation* entails both the identification of the intention of the designer and the appreciation of the film by the viewers.

We address the empirical testing of this model in the following sections.

### 2.3 The think-aloud paradigm

Several studies have been carried out in order to analyze how people make sense of visual metaphors within the genre of advertising (Forceville 1996; Phillips 1997). In their experiments, participants were asked to look at several advertisements containing visual metaphors and then they had to answer several pre-designed questions directed to draw their meanings from those ads in a written report. Other scholars (McQuarrie and Mick 1999) proceeded the same way but their participants had to respond verbally to some questions instead of writing their answers in a form.

In general, all the empirical studies whose aim is to discern how people make sense of visual metaphors in ads follow the same pattern; they all “force” participants to answer guided questions. Although these findings provide valuable insights into the way people interpret visual metaphors within the genre of advertising, they also assume some constraints (Šorm and Steen 2013).

There might be a risk of interviewer bias regarding the targeted questions that participants are asked to answer. Instead of looking at the pictures and talking naturally about what comes to their minds, participants’ thoughts navigate from one question to another, which means that their attention and ideas are guided to the focus of each question, even if they have not discerned any metaphorical meaning in the picture. This completely annuls the possibility of a “simple looking” process, a process that would entail just perceptual identification and processing of visual components. Another potential limitation about this type of studies is that “the order in which the questions are asked may influence the order in which thoughts come to the participants’ minds” (Šorm and Steen 2013: 4). This drawback, however, may not be significant if the analyst is not interested in the order of the metaphor processing stages.

If we are to investigate the mental operations of offline metaphor construction in films, another approach should be tackled for the present study to avoid these two main constraints.

Think-aloud protocols are highly used in testing different types of tasks (Pressley and Afflerbach 1995; Van Den Haak et al. 2003; Altuntaç 2015). The method is extensively applied to investigate “people’s cognitive processes during the execution of a wide range of tasks” (Van Den Haak et al. 2003: 339), and instructions on how to manage research using think-aloud experiments is offered in several textbooks (Barnum 2002; Dumas and Redish 1999; Rubin 1994). Even Nielsen (1993: 195) states, “Thinking aloud may be the single most valuable usability engineering method.” Ericsson and Simon’s model (1993) for evoking verbalizations in usability tests is the most widespread technique to collect verbal data and to validate theoretical cognitive models.

Two types of think-aloud protocols are distinguished in the scientific literature: concurrent think-aloud protocols and retrospective think-aloud protocols (Bowers and Snyder 1990; Hoc and Leplat 1983; Nielsen 1993; Van Den Haak et al. 2003). The former consists in participants doing the task and verbalizing their thoughts at the same time, whereas in the latter, participants are asked first to do the task, and then verbalize their thoughts after the task completion. Both are described as equal options for usability tests and data collection (Nielsen 1993), and although they produce similar results in task completion and task performance (Hoc and Leplat 1983), some research found that the verbalizations of the retrospective think-aloud tasks consisted more in explanations and suggestions



whereas the ones produced in the concurrent think-aloud tasks resulted more in simple descriptions (Bowers and Snyder 1990). In an empirical study in which concurrent and retrospective think aloud designs are compared for a usability test of an online library catalog, Van Den Haak et al. (2003) demonstrate that concurrent and retrospective think-aloud protocols reveal comparable sets of usability problems, which however emerge in different ways: while in retrospective think-aloud protocols, more problems were detected by means of verbalization, in concurrent think-aloud protocols, more problems were detected by means of observation. This suggests that both protocols produce similar results “in terms of quantitative output (but) they differed significantly as to how this output was established” (Van Den Haak et al. 2003: 349). As suggested further in a more recent study, concurrent think-aloud protocols tend to result in verbalized thoughts about the description of the task, whereas retrospective think-aloud protocols often result in “verbalized thoughts about their cognitive operations” (Altuntaç 2015: 7).

The findings of these studies were crucial for the selection of the protocol for our study. Since films are complex materials that demand high levels of attention and focus, and as the understanding and construction of a metaphor may imply complex cognitive processes such as inferencing and deducting, we decided to design a retrospective think-aloud task to collect the data for our particular research.<sup>4</sup>

## 3 Method

### 3.1 Materials

Two TV commercials from different perfume brands were used to collect the data: one for practice in which participants emulated the process of the task, and another one to collect experimental data with the retrospective think-aloud task.

---

<sup>4</sup> The choice for a retrospective paradigm was supported by an initial empirical exploration in which we showed informally the filmic advert to three colleagues and asked them to verbalize aloud their thoughts while they were watching the video. We received extremely negative feedback from all of them: the task was too hard; they could not focus on what they were watching and talk at the same time; they found the task rather annoying because it was too difficult. This finding is in line with Van Den Haak et al. (2003), who found that in the concurrent think-aloud protocols the requirement to think aloud while working had a negative effect on the task performance: concurrent think aloud tasks for complex stimuli is counterproductive. As these authors suggest, and we agree with them, this raises questions about the reactivity of concurrent think-aloud protocols, especially in the case of high task complexity.

Both advertisements included verbal and nonverbal elements, as it is common in audio-visual messages, and they were previously marked for metaphoricity by two independent annotators with the use of FILMIP (Bort-Mir 2019), a method that leads analysts to identify metaphorically-used components in films on a seven-step basis.

Among all the filmic genres, our study was conducted with TV commercials because they contain filmic elements that are specifically designed and combined together in a clip to make the audience change their perspective towards a brand or a product. Metaphors are used by advertisers because they allow the audience to change their perspective; they help them alter their point of view about what is being advertised. As argued by several scholars, “metaphors can be fundamental for persuasion” (O’Shaughnessy and O’Shaughnessy 2003:30). Thus, and also because they are very short clips that last only a few seconds, TV commercials are considered ideal materials for the study of metaphor within the filmic medium.

The advertisement used for practice (see Figure 2) was a TV commercial of the perfume Ricci Ricci (Nina Ricci 2009).

The campaign, released on September 10, 2009, was created by the advertising agency Mazarine Mille Noï (France). In the ad, there is a young woman elegantly dressed, hiding playfully from an attractive young man at the roof of some buildings in Paris. This filmic advertisement contained the following communicative modes: written discourse, music, and visuals.

The second commercial (Adolfo Domínguez 2015), used for the retrospective think-aloud task, was another filmic advertisement from the perfume brand Adolfo



**Figure 2:** Screenshot from *Ricci Ricci's* commercial.

Domínguez, and it was released in 2015 in Spain by the agency China. This filmic ad (see Figure 3) shows a woman fishing roses with a net at a calm sea full of white and pink roses. It contains the following communicative modes: written discourse, spoken discourse, music, and visuals.

The stimuli were presented on the screen of a laptop through the video-sharing online platform YouTube (Ricci Ricci's commercial retrieved from <https://youtu.be/DFC1JUNxelg>, and Agua Fresca de Rosas' commercial retrieved from <https://youtu.be/K2rjjhloL8>).

### 3.2 Participants

We ran the experiment with participants from three different countries: Spain, USA, and Iran. We chose these three countries because they appear to be quite different from one another in terms of cultural beliefs, habits and conceptualizations in general, as well as geographically distant. We therefore speculated that the differences in performing the think-aloud task would emerge in a clearer way.

1. Spain: 30 Spanish native speakers (24 women and six men) from 20 to 25 years old. They were students at Universitat Jaume I, Castellón, in their third course of the English Studies Degree. They participated voluntarily with the reward of an extra 0.3 points in their final grade.
2. USA: 30 American English native speakers (21 women and nine men) from 20 to 25 years old. They were enrolled in various BA programs at various academic institutions based in the USA. The experiment took place in Siena, Italy, during a summer program at the International Center for Intercultural Exchange. They participated voluntarily.



Figure 3: Screenshot from *Agua Fresca de Rosas*' commercial.

3. Iran: 30 Persian native speakers (25 women and five men<sup>5</sup>) between 21 and 24 years old all enrolled in Law BA programs at Hazrate Masoumeh University, Iran. They participated voluntarily with the reward of an extra 0.5 points in the final grade.

All participants were told that the purpose of the study was to understand how they made sense of filmic materials.

### 3.3 General procedure<sup>6</sup>

Data were collected in December 2017 (Spanish participants), in July 2018 (American participants), and in November 2018 (Persian participants).

The retrospective think-aloud task took place in a quiet classroom at the different institutions. Each participant was tested individually. Participants took part in the test on a voluntary basis and were free to leave if they found the task uncomfortable. The think-aloud data were recorded with digital recorders, saved in mp3 format and then transcribed. A code was associated to the participants in order to preserve their anonymity.

A set of written instructions to be signed for authorization to use their data (see Appendix I) was given to the participants. The instructions<sup>7</sup> were given in English to the Spanish and American participants and in Persian to the Persian participants. However, as the commercial for the data collection (Adolfo Domínguez 2015) contained the lyrics of the song in English, we included in the instructions a written translation of those lyrics in Spanish for the Spanish participants, and a translation in Persian for the Persian participants. The instructions were based on Ericsson and Simon (1993: 376), and Šorm and Steen (2013).

Our instructions informed participants about the purpose of the task (to gain insight into the way people interpret filmic texts), and about the task itself: they were going to be audio-recorded; a test commercial was going to be projected first for warming up, they would have then some time to say whatever came to their minds, and finally they were going to see the commercial five times for the data collection. Five times was considered appropriate as “an extended period of immersion”

---

<sup>5</sup> The female Persian participants were all students from Hazrate Masoumeh University (a same-sex university in which the co-author worked at the time of the data collection). The five men were from personal connections of the third author.

<sup>6</sup> All the materials related to this study are stored on Open Science Framework at the following link: [https://osf.io/7362d/?view\\_only=23f1a2e14d944b00badf0b288a0bda3e](https://osf.io/7362d/?view_only=23f1a2e14d944b00badf0b288a0bda3e).

<sup>7</sup> The instructions were given in writing in English language to Spanish and American participants (as Spanish participants had a C1 level of English), and they were given in writing in Persian to the Persian participants.

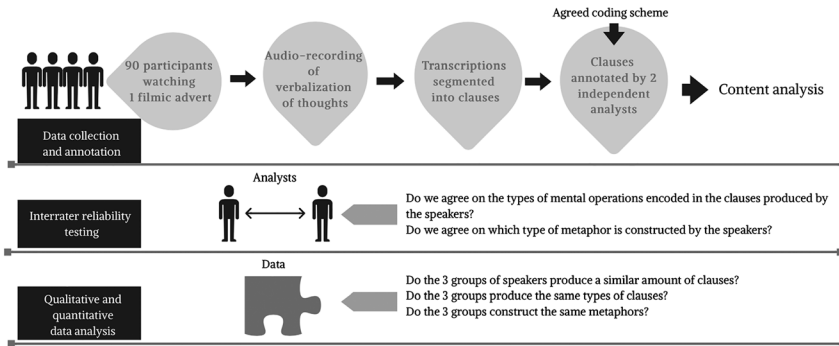


Figure 4: Summary of our empirical research.

(Phillips and McQuarrie 2002: 3). Participants were informed that they would be reminded to keep on talking in the event of a long pause (Ericsson and Simon 1993: 256). Finally, they were given the chance to ask for doubts or clarifications about the instructions or the task. Even though the participants were instructed to verbalize their thoughts at any time, they started speaking only at the end of the fifth projection, which means that we collected just one verbalization per participant.

The practice advertisement was projected to offer participants the opportunity to practice verbalizing their thoughts, as suggested by Ericsson and Simon (1993: 240–241).

Each of the sessions took about 10 min.

The Spanish and American participants verbalized their thoughts in their native language (Spanish and English respectively), and their transcriptions were maintained in their original languages since the two annotators of the content analyses spoke both languages. Even though the Persian participants also verbalized their thoughts in Persian, their transcriptions were translated by the Persian co-author of the paper into English with the aid of Google Translator in order to minimize her intervention in the translations<sup>8</sup>.

### 3.4 Design of the content analyses

All the data (see Figure 4) were segmented into grammatical clauses (Šorm and Steen 2013). A total of 1.342 segments were taken into consideration for the present study.

<sup>8</sup> All the audio recordings are safely stored by the experimenters, in compliance with the privacy rules set by our professional affiliations.

Two independent analysts classified all segments according to their corresponding offline mental operation of filmic metaphor construction. Each segment was only allowed to be coded with just one category.

A codebook was created along the analysis of what the segments really expressed, containing the categories identified for filmic metaphor offline processing, based on the guidelines provided by Bolognesi and colleagues (Bolognesi et al. 2017: 1988) and the measurements used in content analysis described in Krippendorff (2013).

The annotation process included three training sessions through which the annotators formalized and refined the coding scheme, in a series of individual annotations followed by discussions on the disagreements. After three discussion sessions, we developed the final coding scheme where all the improvements were taken into consideration (see Table 1).

Something that should be remarked about category 0 is that incomplete sentences – sentences lacking a subject or predicate – are a frequent feature of spoken language. Therefore, it was expected that the protocols would contain a considerable number of incomplete sentences. If the context in which a segment appeared contained enough clues to infer what the speaker intended to communicate, it was determined what cognitive process corresponded best to that segment. If the context did not help to arrive at the intended meaning, however, the category label “irrelevant” was assigned.

A second content analysis was performed in a second stage, on those clauses that were annotated as Category 2.4 (Metaphor Construction). For this second content analysis, the two authors first scanned the data independently, and agreed on a coding scheme that would include a list of the potential metaphors the participants seemed to be constructing within the film. Then, independently, the two annotators coded the clauses previously coded as “2.4 Metaphor Construction” with one of the seven types of metaphors that constituted the coding scheme, reported in Table 2. Some of these metaphors are clearly related to one another and could constitute mappings of higher, more general conceptual metaphors. For example, *WOMAN IS FISHERMAN* and *MEN ARE FISH* can be linked together under a more general label *LOOKING FOR LOVE IS FISHING*. For the purpose of this specific content analysis, however, we used a simplified coding scheme that involved only agents and patients (nouns) rather than actions and relations, in order to keep the number of categories and their mutual exclusivity manageable. A qualitative analysis of these aspects of the metaphors is provided in the discussion section.

Table 1: Codebook scheme and criteria.

Category label	Criterion	Examples extracted from our data
0		
Irrelevant	When participants talk about not-task related issues, comment on themselves, use interjections, say grammatically incomplete utterances, say something not loud or clear enough to be heard (Šorm and Steen 2013)	<ul style="list-style-type: none"> <li>- <i>Ok, am ...</i></li> <li>- <i>Yeah sorry, am</i></li> <li>- <i>Oh what else</i></li> <li>- <i>That's kind of the vibe I got</i></li> <li>- <i>It has a ...</i></li> </ul>
1. Content description		
1.1 Perceptual elements description	When participants talk about what they see: identification and description of objects, people, music, actions, gestures, colors, spoken and/or written discourse, icons, logos, etc.	<ul style="list-style-type: none"> <li>- <i>And the boat is white</i></li> <li>- <i>And she's like lounging on this boat in the middle of this body of water</i></li> <li>- <i>And the body of water was very calm</i></li> <li>- <i>There are roses in the water</i></li> <li>- <i>Like the opening shot with the ehm the name of the perfume designer I assume</i></li> </ul>
1.2 Genre processing	When participants deal with the type of filmic text	<ul style="list-style-type: none"> <li>- <i>Am, okay so I thought right away that it would be a perfume commercial</i></li> </ul>
1.3 Product/Title processing	When participants deal with the product or service advertised. When participants deal with the title of the ad	<ul style="list-style-type: none"> <li>- <i>It seems like an elegant perfume</i></li> <li>- <i>They're probably alluding to nice fragrances with flowers</i></li> <li>- <i>And that may be the fragrance it's like light but rosey smelling</i></li> </ul>
1.4 Meaning construction	When participants start inferring possible meanings, deriving from the description of the previous steps (1.1, 1.2, 1.3). This step is like a bridge between Category 1 and Category 2	<ul style="list-style-type: none"> <li>- <i>So she's kind of putting feminine spin on a typically masculine, or what is a normally masculine profession</i></li> <li>- <i>So she is kind of living in bliss</i></li> <li>- <i>So I think it's like she's hungry for scent</i></li> </ul>

Table 1: (continued)

Category label	Criterion	Examples extracted from our data
2. Metaphor construction		
2.1 Incongruity identification	When participants express awareness of an inappropriateness within the filmic scene, for example by acknowledging that something is not as it would normally expected to be	<ul style="list-style-type: none"> <li>- [the boat] it's not black, it's not like a gross brown</li> <li>- In that the lyrics don't like fully fit</li> <li>- But then her fishing boat thing is kind of dinky/ which is kind of weird to me</li> </ul>
2.2 Resolution problem identification	When participants express their difficulty or doubt with solving a filmic metaphor	<ul style="list-style-type: none"> <li>- I can't see why what they did works for selling a perfume like this</li> </ul>
2.3 Metaphor recognition	When participants express their metalingual awareness regarding a metaphor as being non-literal or comparative	<ul style="list-style-type: none"> <li>- They're using efficiently metaphor</li> </ul>
2.4 Metaphor construction	When participants interpret and construct a metaphor in terms of language belonging to both the target and/or the source domain	<ul style="list-style-type: none"> <li>- Looking for or hoping to attract the right fish in the sea; the right man from the sea</li> <li>- The roses are like the men</li> <li>- And like that idea that fish being like food</li> </ul>
3. Evaluation		
3.1 Communicative intention detection	When participants identify/describe the intentions of the commercial	<ul style="list-style-type: none"> <li>- If you wear this perfume, you will smell like these flowers that you caught so.</li> <li>- Like if you, if you buy the perfume you're gonna be like this girl.</li> </ul>
3.2 Appreciation of the commercial/product	When participants comment on the positive or negative overall design of the commercial, or on the positive or negative feelings/thoughts of the product/service advertised	<ul style="list-style-type: none"> <li>- So I mean it's a good perfume ad</li> <li>- I can see being successful to certain crowds, definitely has that appeal</li> <li>- Yeah I would say compared to the commercials I've seen for perfume this one at least seems logical</li> </ul>



## 4 Results

### 4.1 Number of clauses in relation to language speakers and types of clauses

Table 3 reports the average number of clauses produced by each of the three groups of speakers, and the related standard deviations. The number of clauses varied tremendously among participants (high standard deviations), with some participants being much more eloquent and talkative than others. Some general trends could be observed between the three groups: while Spanish and Persian participants tended to formulate a similar number of clauses to describe the advertisement, an ANOVA test confirmed that American participants were significantly more talkative than the other two groups ( $F = 11.55$ ,  $df = 2$ ,  $p < 0.001$ ).

Figure 5a and 5b, instead, show the frequency by which the different types of clauses (i.e., the different categories described in the coding scheme) were used overall by the 90 participants (5a) and specifically by the three groups of speakers (5b). These figures show that participants mostly described what they saw within the filmic advertisement, that is, they formulated sentences in which the elements visually represented were described. This category is followed by the categories identifying clauses in which the participants start to construct meaning within the filmic advertisement, and then to construct metaphors.

Conversely, categories describing clauses in which the participants explicitly acknowledge that there is an incongruity within the video, or acknowledges that there is a metaphor within the filmic advertisement, are scarcely produced. This suggests that their metacognitive awareness of the metaphorical constructions within this advertisement may be not well captured by the think-aloud paradigm.

By means of a Chi square test, we then checked whether there were significantly strong relations between specific types of clauses (i.e. specific categories) and specific groups of speakers. Based on the observed and expected frequencies

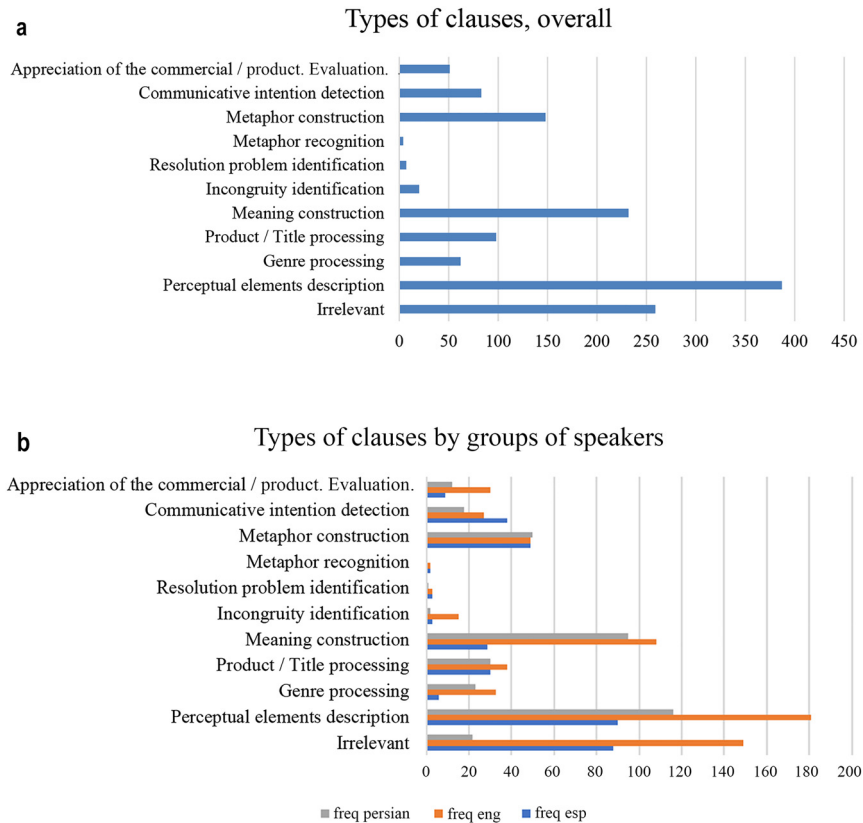
**Table 2:** The coding scheme used to annotate types of metaphors in the second content analysis.

1	PERFUME IS SEA WATER
2	ROSES ARE FISH
3	MEN ARE ROSES
4	ROSES ARE ITEMS SHE WANTS
5	MEN ARE FISH
6	WOMAN IS FISHERMAN
7	NONE OF THE ABOVE (null category)

**Table 3:** Average number of clauses produced by the three groups of participants.

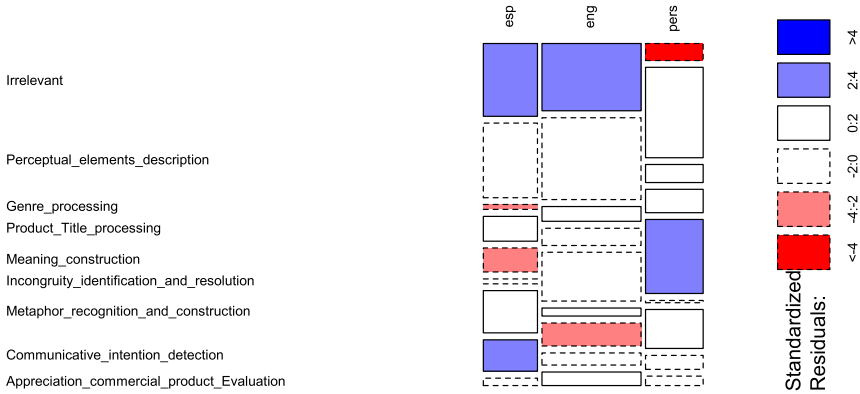
	Average number of clauses produced by participants
Spanish participants	M = 11.3; SD = 6.5
American participants	M = 21.2; SD = 13.2
Persian participants	M = 12.3; SD = 3.8

of production of the various types of clauses, we observed that some clauses were particularly related to specific groups ( $\chi^2 = 128.59, df = 16, p < 0.001$ ). Figure 6 shows the adjusted residuals of this analysis, which suggest that (besides the



**Figure 5:** The types of clauses used by all participants (5a) and by each of the three groups of participants (5b).

### Distribution of categories with adjusted residuals



**Figure 6:** Distribution of categories used by each group of participants (standardized residuals plotted with R statistical computing software, 3.5.1).

category that encompassed irrelevant clauses) the category of Meaning Construction was particularly related to the group of Persian speakers, while the category of clauses labeled as Communicative Intention Detection was particularly related to the group of Spanish speakers. Please note that those categories that displayed a frequency  $< 5$  had to be merged with other sub-categories within the same macro-category type, in order to perform the Chi-square test (observed values lower than five jeopardize the validity of the Chi-square test).

## 4.2 Content analysis: Interrater reliability on the annotation of types of clauses

Two independent annotators annotated the 1.341 clauses produced by the 30 Spanish speakers, the 30 American speakers, and the 30 Persian speakers. The annotators used the coding scheme described in Section 3.4. The results of the content analysis show that the coding scheme, which is enriched with descriptions of the categories and examples, can be reliably applied to the collected data (Krippendorff's  $\alpha = 0.848^9$ ). For the purpose of the interrater reliability test, the

<sup>9</sup> By standard agreement, scores above 0.7 are considered indicators of high reliability. Krippendorff's alpha is preferred to other measurements, such as Cohen's Kappa or Fleiss' Kappa, because of its flexibility (see Bolognesi et al. 2017, for further discussion).

macro and the micro categories were merged into one variable, and the agreement scores were therefore calculated on the complete annotation, that is, a clause annotated as, for example, 24, or 13 (where the first cypher indicates the macro category and the second cypher indicates the nested category).

The disagreements between annotators were then solved in a discussion. The final agreed annotations are publicly released and stored together with all the materials in the OSF online repository at the following link: [https://osf.io/7362d/?view\\_only=23f1a2e14d944b00badf0b288a0bda3e](https://osf.io/7362d/?view_only=23f1a2e14d944b00badf0b288a0bda3e).

### 4.3 Metaphor construction: A cross-cultural perspective

As described in Section 4.2, we then extracted from the codebook containing all the transcribed and annotated protocols (relying on the final agreed annotations) those clauses that were annotated with macro Category 2 (Metaphor Construction) and all its nested categories. On these clauses, we ran additional analyses to annotate these clauses into types of metaphors constructed by the speakers, and to observe the distribution of metaphor types across the three groups of speakers.

Overall, the two annotators coded 136 metaphor-related clauses, using the seven types of metaphor listed in the coding scheme for metaphor analysis, described in Section 3.4. This coding scheme proved to be reliably applied by the two annotators (Krippendorff's Alpha = 0.877).

A first analysis of the metaphor-related clauses showed that, on average, the three groups of participants produced similar numbers of metaphor-related clauses: there are no significant differences between the numbers of metaphor-related clauses produced by Spanish, English and Persian speakers ( $F = 0.003$ ,  $p = 0.9$ ). However, as we indicated above, the English speakers produced on average more clauses than the other two groups, which indicates that, in comparison to the other two groups, English speakers produced a smaller percentage of metaphor-related clauses overall.

As the contingency table displaying the frequencies by which each metaphor was produced (See Table 4) includes empty cells (cells with value 0), we cannot reliably run Chi-square statistics to test the relation between types of metaphors and groups of speakers. Moreover, as the contingency table is larger than a  $2 \times 2$  table, we cannot run a Fisher exact test. Therefore, we analyzed the data in terms of percentages. Figure 7 shows the percentages by which each type of metaphor was produced by each group of speakers.

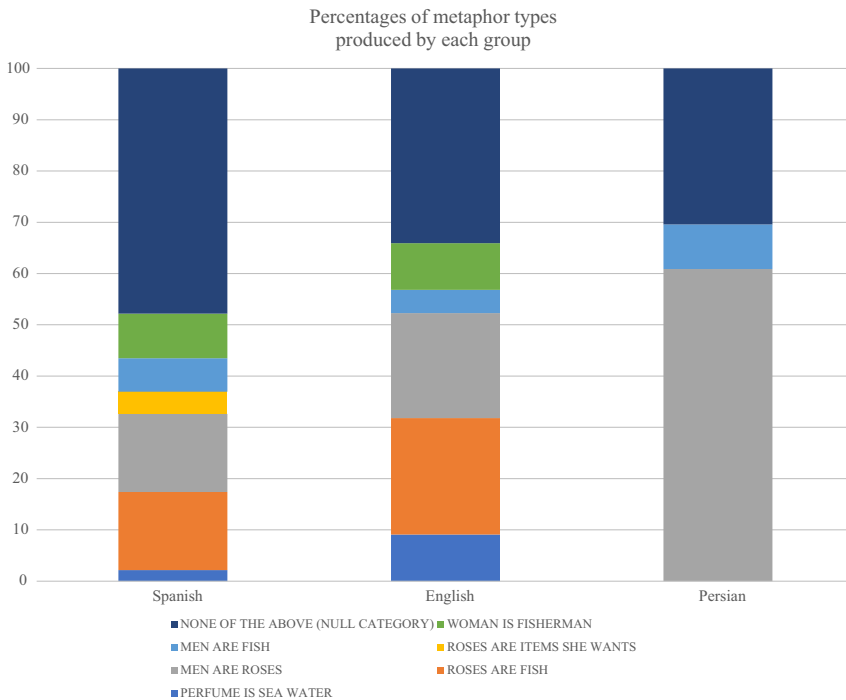
**Table 4:** Frequency by which each metaphor was produced by each group.

	PERFUME IS SEA WATER	ROSES ARE FISH	MEN ARE ROSES	ROSES ARE ITEMS SHE WANTS	MEN ARE FISH	WOMAN IS FISHERMAN	NONE OF THE ABOVE	Total
Spanish	1	7	7	2	3	4	22	46
American	4	10	9	0	2	4	15	44
Persian	0	0	28	0	4	0	14	46
Total	5	17	44	2	9	8	51	136

## 5 Discussion

In the beginning of our study, we formulated the following research questions:

1. Is it possible to determine a list of mental operations that reflect the way in which speakers of different languages describe, analyze and interpret metaphorical filmic advertisements?
2. Are there mental operations that are particularly related to specific groups of speakers?



**Figure 7:** Percentages of metaphor types produced by each group.

3. Do speakers of different languages tend to identify and construct the same metaphors, when exposed to the same metaphorical filmic advertisement?
4. Can these offline cognitive processes be integrated into a theoretical model of filmic metaphor construction?

We ran a series of analyses to address these questions, and we hereby discuss the reported results.

In relation to the first research question, we described the theoretical model provided by Leder et al. (2004) in which they describe the processing of aesthetic experiences and the several distinct cognitive and affective processes that such processing entails, and its adaptation provided by Šorm and Steen (2013) to visual metaphors in still images. We then adapted this theoretical model and the categories included therein to the analysis of metaphors in filmic advertisements. Based on this model, we elaborated a coding scheme in which we annotated the clauses produced by three groups of speakers (Spanish, American, and Persian) in a think-aloud task, in relation to a filmic advertisement of a Spanish brand of perfumes. In Section 4.2, we reported the results of our content analysis and of the interrater reliability test, which show that our coding scheme can be reliably applied to the annotation of the collected data.

In relation to the second research question (Are there mental operations that are particularly related to specific groups of speakers?) in Section 4.1 we reported the analysis of the types of clauses produced by the three groups of speakers. We showed that the category of Meaning Construction was particularly related to the group of Persian speakers, while the category of clauses labeled as Communicative Intention Detection was particularly related to the group of Spanish speakers. Although it is difficult to find an explanation for these differences (they may be due to simple reasons such as personal traits or different academic backgrounds), we can argue that meaning construction relies on semantic knowledge, while detecting the communicative intention is a metacognitive process that relies on pragmatic and communicative knowledge. In this sense, Persian speakers tended to focus more on the semantic knowledge to interpret the filmic advertisement, while Spanish speakers tended to use metacognitive strategies to infer the communicative intentions. On the one hand, this can be explained by the fact that the advertisement has been originally aired in Spain. For the Spanish participants, therefore, grasping the communicative intentions could have been easier, while for the Persian speakers it could have been harder, and this could explain why the latter group focused on different types of knowledge to analyze and interpret this message. Further studies can shed light for the reason behind such difference.

In relation to our third research question (Do speakers of different languages tend to identify and construct the same metaphors, when exposed to the same

metaphorical filmic advertisement?), we ran additional analyses on those clauses that were previously annotated with the Category 2 (Metaphor Construction). We extracted these clauses from the data produced by the speakers in the think-aloud task and analyzed them in the following way. First, we annotated these clauses using a set of possible metaphors that was proposed by one of the analysts, who gave a preliminary inspection of the dataset. We then ran a second content analysis, which we reported in Section 4.3, which showed that this list of possible metaphors could be reliably applied to annotate the clauses. Then, we observed the percentages by which each of these metaphors was produced by each language group, and noticed that Spanish participants tended to construct a more varied set of metaphors, although most of the clauses that they produced were referring to the null category. American English participants showed a similar distribution, but constructed more frequently the *PERFUME IS SEA WATER* metaphor. Persian participants, instead, appeared to us much more homogeneous in the type of metaphor they constructed, which tended to see *MEN* as the unique target domain of the possible metaphors constructed within the filmic advertisement even though, quite interestingly, men do not appear at all in the advertisement. Most clauses produced by Persians relate to the metaphor *MEN ARE ROSES*, many clauses convey other types of metaphor-related information (the null category), and some relate to the metaphor *MEN ARE FISH*.

As Hofstede (2011) puts it, Asian societies expose a more collectivist feature of which Persian society is not an exception in comparison to the more individualistic feature of the Western societies, and thus it could be theorized that the Persian participants in this study tended to be more in harmony with each other in their opinions through associating roses to men. Moreover, such connection between roses and men, even though men did not appear in the filmic advertisement, speaks about the significance of the existence of a male in a female's life for having a sweet fragrant life in the Persian context. In addition, seeing men as fish is a common metaphor in the Persian context and it is used frequently in the everyday life of ordinary people: girls talk about displaying fishing nets for capturing boys' attention as well as several other domain-related metaphoric expressions.

On a qualitative basis, we also observed that, while Spanish speakers considered this (Spanish) advertisement appropriate, both a few Persian and American native speakers indicated that the lady was excessively sexy and this could have been seen as an objectification of the female body for commercial purposes.

From our observations we therefore conclude that, even acknowledging our limitations regarding the small number of participants or the few materials employed for our research, filmic advertisement in a globalized society may be perceived differently by speakers of different cultures, and interpreted in slightly

different ways, constructing different types of metaphors, and evoking different emotions. Therefore, a cognitive model of filmic metaphor construction (our fourth research question) can possibly be proposed, but it has to be flexible enough to accommodate the cultural differences that speakers of different languages display when they interpret metaphorical ads.

## 6 Conclusion

The general aim of this study was twofold: (i) to empirically validate our model of filmic metaphor construction in TV commercials through several analyses and reliability tests, and (ii) to test whether there is a difference in the identification and construction of filmic metaphors across cultures.

Regarding the first aim, it seems that the think-aloud data obtained in the study justify the model, corroborating the existence of the offline mental operations described in Table 1 (Section 3.4) that occur in the minds of the viewers.

Participants devoted several thoughts to the description of the content in the commercials (Category 1: Content Description). This content was described along several levels of granularity: describing the perceptual elements of the videos, talking about the genre, talking about the title or the product advertised, and finally starting the process of meaning construction.

Another mental process that we envisaged in the model and that can be extracted from the think-aloud data is the Metaphor Construction process (Category 2). This is the process where participants identify incongruous elements in the commercials, try to resolve those incongruities, clearly specify that they have identified a metaphor, and they also reconstruct the metaphor.

Finally, our Category 3 (Evaluation) is also present in the verbalized thoughts of the participants, in which they express their feelings about the commercial, the brand or the product, and in which they identify the communicative intention of the ad.

As for the second aim, the analyses performed in our study show that there are certain differences across cultures both in the number of clauses produced and in the types of used categories. Future research could be focusing on whether these differences are not only culture specific but also gender specific. More materials (TV commercials) and a higher number of participants would be needed to carry on with such an investigation. We believe that this first contribution provides a valuable exploration of the cultural differences observed between Spanish, American English, and Persian speakers who were exposed to the same multi-modal commercial message. Given the current political tensions between countries in which the languages hereby analyzed are spoken, we hope to have



contributed in a positive way to show that there are indeed differences in the way in which the same message is processed and interpreted by these different populations. Knowledge about these differences is key for taking wise decisions. Even in the case of commercial advertisements (let aside political communication), deeper knowledge of these differences, together with a greater sensitivity toward diversity, must become core aspects in intercultural communication settings.

## Appendix

### Guidelines for participants:

You are going to watch a TV commercial. You are asked to say aloud what you see and your interpretation of the commercial. There is no wrong answer, so feel free to say whatever comes to your mind at any moment.

What you say is going to be recorded anonymously.

The results of this study are intended to gain insight into the way people process filmic texts.

First you will see a 'test' commercial just as training, and you can say whatever comes to your mind about it.

After this test, you will see the chosen commercial for this experiment five times, and you can start talking (what you say will be recorded) at any moment.

The only thing that I can say to you while you are talking is "Keep on talking," so please don't ask me anything directly.

Do you have any questions?

**Note:** The lyrics of the commercial have been written down to aid with its comprehension.

**Lyrics:** *The ghost who walks/she's on the prowl/for the man she loved/he cut her down.*

## References

- Adolfo Domínguez [Eugenia Silva]. 2015. Adolfo Domínguez: Agua Fresca de Rosas [video file]. <https://www.youtube.com/watch?v=K2rjjhllol8> (Retrieved 24 November 2015).
- Altuntaç, Pelin. 2015. *The comparison of concurrent and retrospective think aloud methods in unmoderated remote usability testing* Leiden: University of Leiden. MA Thesis.
- Barnum, Carol M. 2002. *Usability testing and research*. New York: Longman.

- Bolognesi, Marianna, Roosmaryn Pilgram & Romy van den Heerik. 2017. Reliability in content analysis: The case of semantic feature norms classification. *Behavior Research Methods* 49(6). 1984–2001.
- Bort-Mir, Lorena. 2019. *Developing, applying and testing FILMIP: The filmic metaphor identification procedure*. Castellón: Universitat Jaume I PhD dissertation.
- Bowers, Victoria A. & Harry L. Snyder. 1990. Concurrent versus retrospective verbal protocols for comparing window usability. In *Human Factors Society 34th Meeting, 8–12. October 1990*, 1270–1274. Santa Monica: HFES.
- Coëgnarts, Maarten & Peter Kravanja. 2012. From thought to modality: A theoretical framework for analysing structural-conceptual metaphors and image metaphors in film. *Image & Narrative* 13(1). 96–113.
- Dumas, Joseph S. & Joseph C. Redish. 1999. *A practical guide to usability testing*, Revised edn. Exeter: Intellect.
- Ericsson, K. Anders & Herbert A. Simon. 1993. *Protocol analysis: Verbal reports as data*, Revised edn. Cambridge, MA: MIT Press.
- Forceville, Charles & Eduardo Uriós-Aparisi. 2009. *Multimodal metaphor*. Berlin & New York: Mouton de Gruyter.
- Forceville, Charles. 1996. *Pictorial metaphor in advertising*. London & New York: Routledge.
- Foss, Sonja K. 2005. Theory of visual rhetoric. In Sandra Moriarty, Kent Smith, Gretchen Barbatsis & Keith Kenney (eds.), *Handbook of visual communication: Theory, methods, and media*, 141–152, New Jersey: LEA.
- Goldstein, Thalia R. 2009. The pleasure of unadulterated sadness: Experiencing sorrow in fiction, nonfiction, and “in person.” *Psychology of Aesthetics, Creativity, and the Arts* 3. 232–237.
- Grodal, Torben. 2009. Film aesthetics and the embodied brain. In M. Skovand & O. Vartanian (eds.), *Neuroaesthetics*, 249–260. Amityville, NY: Baywood.
- Hilscher, Michelle C, Gerald C. Cupchik & Garry Leonard. 2008. Melodrama and film noir on today’s big screen: How modern audiences experience yesterday’s classics. *Psychology of Aesthetics, Creativity, and the Arts* 2. 203–212.
- Hoc, Jean M. & Jacques Leplat. 1983. Evaluation of different modalities of verbalization in a sorting task. *International Journal of Man-Machine Studies* 18. 283–306.
- Hofstede, Geert. 2011. Dimensionalizing cultures: The Hofstede model in context. *Online Readings in Psychology and Culture* 2(1). 8.
- Kappelhoff, Hermann & Cornelia Müller. 2011. Embodied meaning construction: Multimodal metaphor and expressive movement in speech, gesture, and feature film. *Metaphor and the Social World* 1(2). 121–153.
- Krippendorff, Klaus. 2013. *Content analysis: An introduction to its methodology*, 3rd edn. Thousand Oaks, CA: Sage.
- Leder, Helmut, Benno Belke, Andries Oeberst & Dorothee Augustin. 2004. A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology* 95(4). 489–508.
- Marković, Slobodan. 2012. Components of aesthetic experience: Aesthetic fascination, aesthetic appraisal, and aesthetic emotion. *i-Perception* 3(1). 1–17.
- McQuarrie, Edward F. & David G. Mick. 1999. Visual rhetoric in advertising: Text-interpretive, experimental, and reader-response analyses. *Journal of Consumer Research* 26. 37–54.
- Nielsen, Jakob. 1993. *Usability engineering*. Cambridge, MA: AP Professional.
- Nina Ricci [cozenimus]. 2009. “Ricci Ricci” perfume commercial with Jessica Stam [video file]. <https://www.youtube.com/watch?v=DFClJUNxelg> (Retrieved 31 October 2009).
- O’Shaughnessy, J. & N. O’Shaughnessy. 2003. *Persuasion in advertising*. London: Routledge.

- Ortiz, María J. 2014. Visual manifestations of primary metaphors through mise-en-scène techniques. *Image & Narrative* 15(1). 5–16.
- Pérez-Sobrino, Paula. 2017. *Multimodal metaphor and metonymy in advertising*, Vol. 2. Philadelphia & Amsterdam: John Benjamins Publishing Company.
- Phillips, Barbara J. & Edward F. McQuarrie. 2002. The development, change, and transformation of rhetorical style in magazine advertisements 1954–1999. *Journal of Advertising* 31(4). 1–13.
- Phillips, Barbara J. & Edward F. McQuarrie. 2004. Beyond visual metaphor: A new typology of visual rhetoric in advertising. *Marketing Theory* 4(1-2). 113–136.
- Phillips, Barbara J. 1997. Thinking into it: Consumer interpretation of complex advertising images. *Journal of Advertising* 26(2). 77–87.
- Pressley, Michael & Peter Afflerbach. 1995. *Verbal protocols of reading: The nature of constructively responsive reading*. Hillsdale, NJ: Erlbaum.
- Richards, Ivor A. 1965. The philosophy of rhetoric. In John Constable (ed.), Vol. 94. New York: Oxford University Press.
- Rubin, Jeffrey. 1994. *Handbook of usability testing: How to plan, design, and conduct effective tests*. New York: Wiley.
- Šorm, Esther & Gerard Steen. 2018. Towards a method for visual metaphor identification. In Gerard Steen (ed.) *Visual metaphor: Structure and process*. 18–47. Amsterdam: John Benjamins Publishing Company.
- Šorm, Esther & Gerard Steen. 2013. Processing visual metaphor: A study in thinking out loud. *Metaphor & the Social World* 3(1). 1–34.
- Steen, Gerard J. 2007. *Finding metaphor in grammar and usage: A methodological analysis*. Philadelphia & Amsterdam: John Benjamins Publishing Company.
- Van Den Haak, Maaïke, Menno De Jong & Peter Jan Schellens. 2003. Retrospective vs. concurrent think-aloud protocols: Testing the usability of an online library catalogue. *Behaviour & Information Technology* 22(5). 339–351.

## Bionotes

### Lorena Bort-Mir

Universitat Jaume I, Castelló de la Plana, Spain

[bortl@uji.es](mailto:bortl@uji.es)

Lorena Bort-Mir holds her Ph.D on Applied Languages, Literature and Translation, and she's is an associate lecturer at Universitat Jaume I, Castellón, within the English Studies Department. She is interested in multimodal metaphor analysis and processing in films. She developed FILMIP, the filmic metaphor identification procedure, in her doctoral thesis. She published articles and papers in various journals, and she was invited as keynote speaker to various international conferences.

### Marianna Bolognesi

University of Bologna, Bologna, Italy

[marianna.bolognesi@gmail.com](mailto:marianna.bolognesi@gmail.com)

Marianna Bolognesi is senior assistant professor in Linguistics at the University of Bologna, Italy. Her focus is on the relation between language and thought, and on the semantic representation of word meaning in mind and its expression in the pictorial and verbal modes of communication. In

2017–2019 she worked at the University of Oxford. In 2015–2017, she was an EU Marie Curie awarded postdoctoral researcher at the Metaphor Lab, University of Amsterdam.

**Susan Ghaffaryan**

Allameh Tabataba'i University, Tehran, Islamic Republic of Iran

**[susan.ghaffaryan@gmail.com](mailto:susan.ghaffaryan@gmail.com)**

Susan Ghaffaryan is an assistant professor in the English Studies Department at AllamehTabataba'i University, Tehran, Iran. Her research interest is in intercultural communication.

## **AUTHORS' CONTRIBUTION STATEMENT:**

**Lorena Bort-Mir, Marianna Bolognesi and Susan Ghaffaryan**

**Cross-cultural interpretation of filmic metaphors: A think-aloud experiment**

**<https://doi.org/10.1515/ip-2020-4001>**

This article is the result a collaboration between the first two authors, LBM and MB. The third author, SG, contributed in the data collection, providing data for the Persian speakers specifically and in the qualitative analysis of the Persian data.

The empirical study was conceived and designed by LBM and MB; Spanish data was collected by LBM, English data by MB and Persian data by SG.

For the specific concerns of the Italian academic attribution system, MB is responsible for writing sections 3.4, 4, 6.

4/11/20  
Bologna, Italy

Signatures:

LBM

A handwritten signature in blue ink, appearing to be 'Lorena Bort-Mir', is written over a circular blue stamp. The stamp contains the text 'UNIVERSITY OF BOLOGNA' and 'LIBRARY'.

MB

A handwritten signature in blue ink that reads 'Marianna Bolognesi'.

SG

A handwritten signature in blue ink that reads 'Susan Ghaffaryan'.