

A Text-based Search Interface for Multimedia Dialectics

Katerina Pastra

Institute for Language & Speech Processing
Athens, Greece

kpastra@ilsp.athena-innovation.gr

Eirini Balta

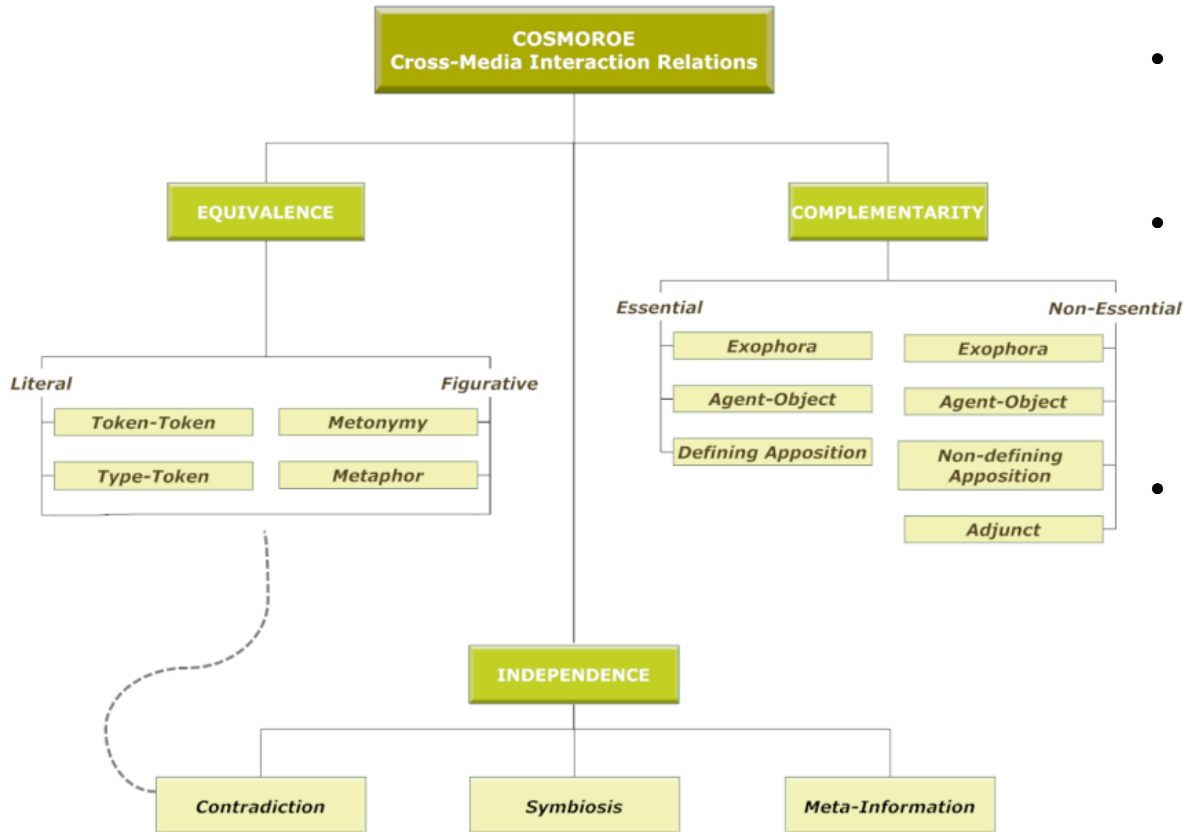
Institute for Language & Speech Processing
Athens, Greece

ebalta@ilsp.athena-innovation.gr

COSMOROE & COSMOROE Search Engine

- COSMOROE (CrOSs-Media inteRactiOn rElations)
 - A framework for modelling multimedia dialectics (semantic interplay between images, language and human actions)
 - Defines a number of semantic relations between different modalities, for formulating multimedia messages
- COSMOROE Search Engine
 - A text-based search engine for multimedia documents
 - A support tool for COSMOROE framework
 - Indexes and retrieves information from audiovisual files, manually annotated
 - Facilitates the active exploration and presentation of the semantic interplay between different modalities
 - Allows interfacing with semantic lexical resources for automatically expanding and generalising the semantic relations

COSMOROE Framework

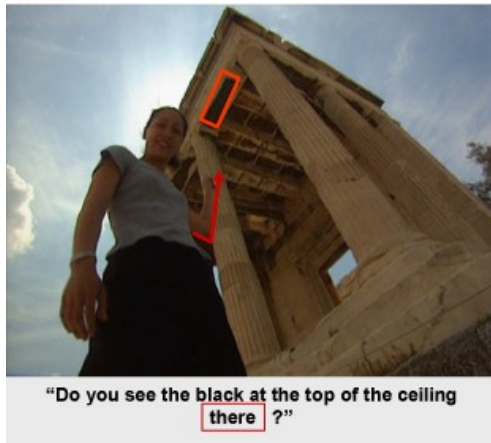


- **Equivalence:** the information expressed by the different media is semantically equivalent (paradigmatic relations between modalities)
- **Complementarity:** the information expressed in one medium is (an essential or not) complement of the information expressed in another medium (syntagmatic relations between modalities)
- **Independence:** each medium carries an independent message, the combination of which creates the multimedia message

COSMOROE Examples

Equivalence

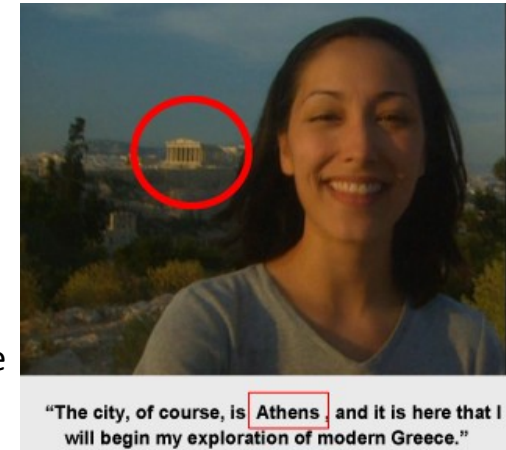
Literal



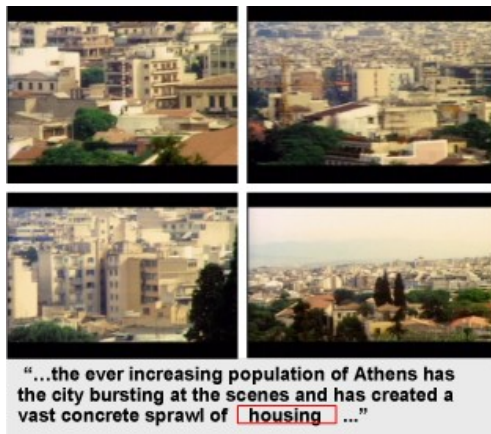
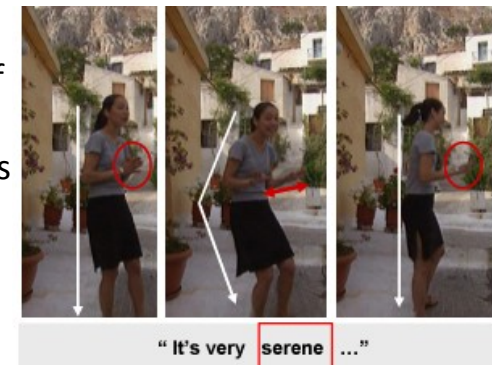
Token-Token: the media refer exactly to the same entity, uniquely identified as such

Metonymy: the modalities refer to different entities but the user considers them semantically equivalent, the referents come from the same domain, have same array of associations, there is no transfer of qualities from one to another

Figurative



Metaphor: the modalities refer to different entities of different domains, the user considers them semantically equivalent, there is a transfer of qualities



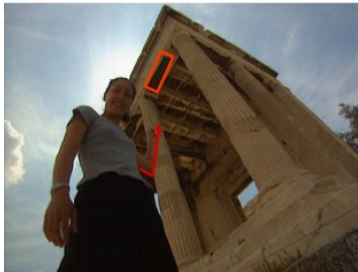
Type-Token: one modality refers to a class of entities and the other to one or more representative members of the class

COSMOROE Examples

Complementarity

Essential

the meaning of what is communicated is clearly comprehended only when information by all participating media is combined



"Do you see the black at the top of the ceiling there?"

Exophora: cases of "anaphora", in which one medium resolves the reference made by another, signals of resolution are present in discourse

Exophora: cases of "anaphora", in which the entity referred to in one medium is revealed in another, though this is not vital for understanding the message



"...The city is a jumble of the ancient and the modern..."

Agent-Object: one modality reveals the agent or the object of the action, state or event expressed by another, but this information is also mentioned in the close context or easily inferred



"...We went shopping in Oxford Street..."

Adjunct: one modality functions as an adjunct to the other (place-position, place-direction, manner)



"...we are heading to Patmos..."

Non Defining Apposition: one medium reveals a generic property/characteristic of the very concrete entity mentioned by another medium, which is not vital for the clear comprehension of the message

Agent-Object: cases of "ellipsis", in which one medium reveals the agent or object of an action, event or state, expressed by another



"...hold, hold..."

Defining Apposition: one medium provides extra information to another, which identifies or describes something or someone (the relation is tied to the specific context and should not be considered as generally valid)



"...Everyone goes to the festival, including the President of Greece..."

COSMOROE Examples

Independence



Contradiction: one modality expresses the exact opposite information to what the other expresses (irony, humour, editorial mistake), one may distinguish all subtypes that Equivalence has

...and of course, **no tourists** come here because...

Symbiosis: cases of different pieces of information being expressed by the media, the conjunction of which (conjunction in time or space) serves "phatic" communication purposes, i.e. one medium provides some information and the other shows something that is thematically related, but does not refer or complement that information in any way



So, what the women were doing then? The women were doing all the jobs, basically...

Meta-Information: one medium reveals extra information on what the other expresses, through its own means of realisation (e.g. camera movement) or through non-propositional aspects of its realisation (e.g. beats, prosody etc.), the medium expresses information that comments on aspects of what the other expresses, going beyond the message communicated to creation-related comments (who created the message, when, why, how – cf. typical archival metadata)

COSMOROE Search Engine

Search for:

Simple Search: targets the general users, queries the audiovisual file in a straightforward manner, like in any text-based search engine

The engine allows one to:

- perform a simple text-based search, taking advantage of the COSMOROE relations -behind the scenes- for more precise and intelligent retrieval within multimedia archives, and
- perform an advanced text-based search, filtering the query with criteria that are directly related to multimedia semantics and to modality specific information

Search is done:

- using single or multiword query terms (keywords) that are mentioned in the transcribed speech (or other text) of the video or in the visual labels set of its visual-units, in order to find instantiations of different semantic relations in which they participate
- using a pair of single or multiword query terms (keywords) that are related through (un)specified semantic relations
- for specific types of relations and find out how these are realized through actual instances in a certain multimedia context
- for specific modality types (e.g. specific types of gestures, image-regions etc.) and find out all the different relations in which they appear

Advanced Search: is addressed to people with a special interest in multimedia semantics and/or multimedia system development, multimedia relations act as semantic filters

Search for:

Query1	Operator	Query2
<input type="text"/>	<input type="radio"/> OR	<input type="text"/>
Query1 Type	<input checked="" type="radio"/> AND	Query2 Type
Visual Label Transcribed Text Multiple selection		Visual Label Transcribed Text Multiple selection
Query1 Type Attributes		Query2 Type Attributes
Gesture Type Shot Type Multiple selection		Gesture Type Shot Type Multiple selection
<input type="button" value="Clear"/>		<input type="button" value="Clear"/>

Search by Multimedia Relation Type:

Relation Type Level	Relation Types
Type First Level Subtype Second Level Subtype Third Level Subtype Multiple selection	<input type="text"/>
<input type="button" value="Clear"/>	










Search by Modality Type:

Modality Type	Modality Attributes
Body Movements Gestures Shots Keyframe Regions Multiple selection	<input type="text"/>
<input type="button" value="Clear"/>	

COSMOROE Search Engine Results

Results

DISPLAYING RESULTS 3-9 OF 9 FOUND

Content	Transcribed Text	Relation File
	Although they are now scarce, it is still possible to find natural sponges in the waters around Santorini.	View Relation Detail
	I am trying to ah trim it, to clean the bottom of it because they are freshly now cut.	View Relation Detail
	What are you doing to that?	View Relation Detail
	Can I see that.	View Relation Detail
	You can..You can lift a Rolls Roils with this one!	View Relation Detail
	Okay? This one here first of all it's so strong you cannot tare it.	View Relation Detail
	These are the local ones, not the import ones.	View Relation Detail
	What are you doing to that?	View Relation Detail
	Where did this come from?	View Relation Detail

[Save as HTML](#) [Save as Text](#)

Results

DISPLAYING RESULTS 1-20 OF 28 FOUND
RESULTS: 1-20 [Next](#)

Relation File	Transcribed Text
GT-Greece-B-EN Adjunct_006.xml	Get there early. Particularly if you are going to a popular island.
GT-Greece-B-EN Adjunct_000.xml	Now, I am heading to Patmos and it's a pretty long journey.
GT-Greece-B-EN Adjunct_002.xml	Most people visit Patmos just as a day-trip, they visit the monastery and get back on their boat, so it actually has a laid back, sleepy town feel.
GT-Greece-B-EN Adjunct_003.xml	Octopus is on just about every menu, cooked in just about every way you can imagine, so I thought I would go out and see how they catch it and with Sozos and he kind off does it in a very traditional way, I mean let's see how he catches those suckers!
GT-Greece-B-EN Adjunct_007.xml	Mykonos is kind off like the San Trope, I guess of the islands. This is a big holiday weekend
GT-Greece-B-EN Adjunct_012.xml	And speaking of leaving, I've got a boat to catch at seven o'clock to take me to Crete.
GT-Greece-B-EN Adjunct_013.xml	Wo!Too early in the morning for me! But I got my ticket, I am headed to Crete...only about a three hour ferry ride. That's not bad.
GT-Greece-B-EN Adjunct_014.xml	Crete is a very popular destination with over two million tourists visiting each year.
GT-Greece-B-EN Equivalence_034.xml	GREEK ISLANDS
GT-Greece-B-EN Equivalence_092.xml	Yay! There is my boat! Okay!
GT-Greece-B-EN Equivalence_109.xml	Most people visit Patmos just as a day-trip, they visit the monastery and get back on their boat, so it actually has a laid back, sleepy town feel.
GT-Greece-B-EN Equivalence_181.xml	Mykonos is kind off like the San Trope, I guess of the islands. This is a big holiday weekend
GT-Greece-B-EN Equivalence_232.xml	To see some really ancient Greek sites, you can take the twenty minute boat trip from Mykonos town to nearby Dilos.
GT-Greece-B-EN Equivalence_317.xml	And speaking of leaving, I've got a boat to catch at seven o'clock to take me to Crete.
GT-Greece-B-EN Equivalence_318.xml	And speaking of leaving, I've got a boat to catch at seven o'clock to take me to Crete.
GT-Greece-B-EN Equivalence_319.xml	DIMITROULA
GT-Greece-B-EN Equivalence_321.xml	Wo!Too early in the morning for me! But I got my ticket, I am headed to Crete...only about a three hour ferry ride. That's not bad.
GT-Greece-B-EN Equivalence_437.xml	SOFIA
GT-Greece-B-EN Equivalence_184.xml	It's a six hour ferry journey from Pireas to Mykonos and costs about five thousand drachma, which roughly fourteen dollars.
GT-Greece-B-EN Equivalence_183.xml	It's a six hour ferry journey from Pireas to Mykonos and costs about five thousand drachma, which roughly fourteen dollars.

[Save as HTML](#) [Save as Text](#)

- The search results are the multimedia files that matched the given query
- They are presented in a list, accompanied by the transcribed text
- Simple Search results are centered around the video clips presenting the relation files, while Advanced Search results are centered around the relation files themselves
- Simple Search results are sorted according to a modified tf/idf score, tuned to favour equivalence and apposition relations, over contradiction, complementarity and symbiosis relations
- The option of saving the content of the retrieved relation files is also provided
- Statistics for the various types and subtypes of relations, as well as for the modalities participating in relations are provided

COSMOROE Search Engine Relation File

Field	Value
Video	GT-Greece-B-EN.avi
Topic	Travel in Greece: Santorini
Transcript	
Overlapping Utterance	These are the local ones, not the import ones.
Anchor Text	
Anchor Overlapping Utterance	these
Anchor Overlapping Utterance Type	token
Keyframe Regions	
Keyframe	
Keyframe Label	<i>sponge</i>
Keyframe Reference ID	51663
Relation	
Relation File ID	relations/GT-Greece-B-EN_Εκοφωρα_117.xml
Relation Type	complementarity
Relation First Level Subtype	essential
Relation Second Level Subtype	εκοφωρα
Relation Video	View video

Save as HTML

Save as Text

Close

- Individual retrieved relation files are accessible, in which the complete structure of the multimedia relation is presented, exposing the semantic interplay between different modalities
- Depicted here are textual and visual components of the relation, information and comments about the type, as well as highlighted indications about the query matches. The user can also watch the video clip "containing" the whole relation, or the video clips of the modalities that participate in it and see static images of specific objects, with the objects' contour highlighted
- Each individual relation can also be saved