

# MULTIMODALITY IN CONVERSATION ANALYSIS: A CASE OF GREEK TV INTERVIEWS

Maria Koutsombogera<sup>1,2</sup>

Lida Touribaba<sup>2</sup>

Harris Papageorgiou<sup>1</sup>

<sup>1</sup>Institute for Language and Speech Processing (ILSP), Athens, Greece

<sup>2</sup>University of Athens, Department of Linguistics

# Multimodality in TV conversations

2

- Semantic & Pragmatic relations between distinct modalities in TV discussions
- Communication through words, prosody, body posture, hand gestures, gaze patterns, facial expressions etc.
- Constraints of TV settings
- Motivation
  - ▣ Identification and interpretation of non-verbal expressions
  - ▣ Message formation and modification
  - ▣ Situational and conversational context (Heritage, 2005; Ilie, 2001)
  - ▣ Multimodal Corpus development

# Methodology

3

- Data collection
- Speech transcription (**Transcriber**)
- Video annotation (**ELAN**)
- Analysis & interpretation

# Corpus

4

- 66 minutes
- One-to-one interactional pattern
- No audience
- Question-answer sequences
  
- Interview 1: institutional
- Interview 2: casual
- Interview 3: semi-institutional



I1



I2



I3

# Multimodal Annotation

5

- Description and analysis of multimodal signals (gestures, facial displays, body posture, *speech*)
- Annotation scheme (*MUMIN- Allwood et al., 2005*)
- 2 levels of representation
  - ▣ Description
    - *Coarse* annotation of *selected* non-verbal expressions
    - Semiotic types
  - ▣ Communicative (semantic/pragmatic) functions
    - Feedback, turn management (*Sacks et al., 1974*)
    - Multimodal relations (*Poggi & Magno Caldognetto ,1996*)

# Annotation scheme

6

Descriptive features	Facial Displays	<b>General</b> [smile, laughter, scowl] <b>Gaze</b> [down, sideways,...] <b>Eyes</b> [semi-closed, wide open,...] <b>Lips</b> [protruded, retracted]	<b>Eyebrows</b> [frown, raise] <b>Mouth</b> [open, closed] <b>Head</b> [nod, jerk,...]
	Gestures	<b>Handedness</b> [single, both] <b>Trajectory</b> [up, down, sideways, complex]	
	Body Posture	<b>Torso movement</b> [forward, backward, left, right]	
	Semiotic Type	Deictic Non-deictic Iconic Symbolic	
Communicative functions	Feedback	<b>Give</b> <b>Elicit</b>	[perception/understanding, acceptance, emotions/attitudes]
	Turn Management	<b>Turn gain</b> [take, accept] <b>Turn end</b> [yield, elicit, complete]	
	Multimodal Relations	Repetition Addition Substitution Contradiction Neutral	



Grid	Text	Subtitles	Controls	
S1-Semiotic Type				
Nr	Annotation	Begin Time	End Time	Duration
1	N-D	00:00:00.220	00:00:06.700	00:00:06.480
2	N-D	00:00:06.700	00:00:08.390	00:00:01.690
3	N-D	00:00:08.390	00:00:12.470	00:00:04.080
4	S	00:00:28.320	00:00:29.470	00:00:01.150
5	N-D	00:00:29.470	00:00:31.220	00:00:01.750
6	N-D	00:00:34.590	00:00:35.730	00:00:01.140
7	N-D	00:00:35.730	00:00:36.740	00:00:01.010
8	N-D	00:00:37.020	00:00:37.890	00:00:00.870
9	N-D	00:00:58.400	00:00:59.653	00:00:01.253
10	N-D	00:01:08.923	00:01:10.653	00:00:01.730
11	N-D	00:01:12.533	00:01:18.393	00:00:05.860
12	N-D	00:01:43.120	00:01:44.660	00:00:01.540
13	N-D	00:01:44.660	00:01:44.663	00:00:00.003
14	D	00:01:52.140	00:01:53.270	00:00:01.130

00:00:29.260

Selection: 00:00:00.000 - 00:00:00.000 0

Media control buttons: Play, Stop, Previous, Next, Full Screen, Repeat, Shuffle, Selection Mode, Loop Mode.

	00:00:29.000	00:00:30.000	00:00:31.000	00:00:32.000	00:00:33.000	00:00:34.000	00:00:35.000	00:00:36.000	00:00:37.000	00:00:38.000
<b>S1-Gestures</b> [14]	SH_U	SH_R_U_D					SH_sideways			
S1-Facial expressi [22]	Eyebrows_R	Eyebrows_R					Head_Side_Turn	Head_Side_Turn		Eyebrows_R
S1-Semiotic Type [21]	S	N-D					N-D	N-D		N-D
S1-Feedback [23]	FG_non_A	FG_non_A					FG_non_A	FG_non_A		FE_A
S1-Turn Managem [20]	TM_take	TM_hold					TM_take	TM_hold		TM_hold
S1-Multimodal Rel [25]	A	A					A	A		A
S2-Gestures [18]					SH_U					
S2-Facial expressi [30]					Eyebrows_R			Gaze_Towards		
S2-Semiotic Type [32]					N-D			N-D		
S2-Feedback [31]					FG_A			FE_non_A		
S2-Turn Managem [28]					TM_take			TM_take		
S2-Multimodal Rel [36]					A			A		
S1-Speech [26]	άλλο αυτό λοιπόν	απαντάτε τώρα στον [κύριο*-]	θέλω να κάνω θέλω άλλο να κάνω				ενώ μελετούσα να κάνω κάτι άλλο όμως			
S2-Speech [39]					βεβαίως έχει λάθος από [τηνFr			από την αφητηρία		όχι για πολιτικά

# Semiotic types (1 / 2)

8

## Deictic



## Non-deictic





# Semiotic types (2/2)

9

## Iconic



## Symbolic



# Feedback (1 / 2)

10

Accept



Non-accept



Perception/Understanding



# Feedback (2/2)

11

## □ Attitudes/ Emotions



# Turn Management

12

□ Turn take



□ Turn hold



□ Turn complete



□ Turn elicit



## Distribution of semiotic and communicative features over interview types

	Value	I1	I2	I3
Semiotic Types	Deictic	5.2%	6.9%	<b>7.4%</b>
	Non Deictic	<b>88.1%</b>	73.1%	75.9%
	Iconic	4.6%	<b>15.9%</b>	13.8%
	Symbolic	2.1%	<b>4.1%</b>	2.9%
Multimodal Relations	Repetition	15.6%	<b>17.8%</b>	6.8%
	Addition	51.7%	54.6%	<b>71.4%</b>
	Contradiction	<b>1.7%</b>	0.4%	0.3%
	Substitution	19.8%	19.9%	<b>20.2%</b>
	Neutral	<b>11.2%</b>	7.3%	1.3%
Turn Management	Turn Take	<b>44.8%</b>	12.1%	16.4%
	Turn Yield	1.9%	8.2%	<b>9.2%</b>
	Turn Accept	2.7%	<b>8.8%</b>	7.4%
	Turn Hold	42.1%	<b>45.2%</b>	40.3%
	Turn Elicit	5.4%	<b>15.3%</b>	14.2%
	Turn Complete	3.1%	10.4%	<b>12.5%</b>
Feedback (Emotions/Attitudes)	Positive	4.4%	12.1%	<b>27.3%</b>
	Negative	<b>9.7%</b>	1.6%	1.9%

# Future steps & exploitation

14

- Add prosodic features
- Enrich corpus with more interview types
- Concrete application scenarios
  - Summarization
  - Indexing&retrieval
  - Avatars
- Intercultural study



Thank you!