

IM2 Phase III

IP3: Social Signal Processing

IP Head: A.Vinciarelli (Idiap)

Deputy IP Head: F.Valente (Idiap)

IM2 Site Visit,
November 12-13, 2008



FONDS NATIONAL SUISSE
SCHWEIZERISCHER NATIONALFONDS
FONDO NAZIONALE SVIZZERO
SWISS NATIONAL SCIENCE FOUNDATION



Outline

- Part I – **Effectiveness of Delivery**
 - Personality and effectiveness of delivery
 - Automatic personality perception from speech
- Part II – **Engagement in group interactions**
 - Engagement and preference structures
 - Speaker diarization and preference structures
- Part III – **Networking**
 - Liaison programs (SSPNet and Affective Sciences)
 - Dissemination events

Part I

- Effectiveness of Delivery

Personality

“[Personality is the latent construct accounting for] *individuals' characteristic patterns of thought, emotion, and behavior together with the psychological mechanisms – hidden or not - behind those patterns*”

D.C. Funder, “*Personality*”, Annual Review of Psychology, 52:197-221, 2001.

Personality Perception

- Experiments show that we **attribute personality traits** in less than 100 ms
- **Speech** is the modality that influences most the attribution of **traits related to competence**
- The process **is not necessarily accurate**, but it drives our behavior with respect to others

S.Uleman, S.Saribar, C.Gonzalez, “*Spontaneous inferences, Implicit impressions, and implicit theories*”, Annual Reviews Of Psychology, 59:329-360, 2008.

Personality and the “Big Five”

- **Extraversion**: active, assertive, energetic
- **Agreeableness**: generous, kind, sympathetic
- **Conscientiousness**: efficient, competent
- **Neuroticism**: anxious, tense, touchy, unstable
- **Openness**: curious, insightful, original, artistic

G.Saucier, L.Goldberg, “*The language of personality: lexical perspectives on the five-factor model*”, in J.S.Wiggins (ed.), “The Five-Factor Model of Personality”, 1996.

Measuring the “Big Five”

- The Big Five are “measured” via questionnaires including a large number of items (40 to 60)
- This work uses the **BFI-10**, short version of the Big Five Inventory
- The measurement consists of five scores in the range [-4,4], one per trait

B.Rammstedt and O.John, “*Measuring personality in one minute or less*”, Journal of Research in Personality, 41:203-212, 2007.

The Data

Total Number of clips	640
Total Number of identities	330
Total duration	7 hours
Average duration	40 sec.
Clips with professional speakers	309
Clips with non-professional speakers	331

Judges assess **10 second long** extracts that do not include proper names, English words or emotional speech

Personality Assessment

Total number of judges	3
Mother tongue	Parsi
Clip ordering	random
Clips per judge	640
Inter-Judge Agreement	0.3
Clips with IJ Agreement > 0.5	279

Judges do not understand the content of the clips, they are driven by purely nonverbal aspects

Preliminary Results

Trait	All	Hi Agr.
Extraversion	76.3%	79.4%
Agreeableness	63.0%	64.7%
Conscientiousness	72.0%	75.7%
Neuroticism	63.0%	63.6%
Openness	57.9%*	62.8%

- Two classes (above average and below average)
- Features are entropy of main prosodic characteristics (pitch, energy, rhythm, formants)
- Classification performed with SVM (K-fold validation)

Part II

- Engagement in Group Interactions

Preference Structures

“If one fails to deny an accusation, a denial is noticeably absent and is cause for inference, the most common inference being that the accusation is true [...] When someone makes an attribution about you, contradict, unless you want others to understand that you accept the truth of the attribution”

J.Bilmes, “*The Concept of Preference in Conversation Analysis*”,
Language in Society, 17(2):161-181, 1988

Preference Structures and Engagement

- In presence of conflicts, contradiction is an evidence of engagement
- If sufficiently robust, the evidence is expected to allow one to:
 - Reconstruct the **structure of the conflict**, i.e. who (dis-)agrees with whom and who moderates
 - Use of the structure **as an a-priori information** in speaker diarization

The Canal 9 Corpus

Total Number of Debates	45
Participants per debate	5
Total Duration	28 hours

Audio annotations :

Speaker labels and speaking time (who speaks when)

Group interaction annotations

Roles: moderator and guest speakers

Conflict structure: coalitions composition

The Approach

Class	Moderator	Group 1	Group 2
Moderator	0%	51%	49%
Group 1	68%	7%	26%
Group 2	67%	25%	8%

- The goal is to find the mapping between participants and classes
- The mapping is the one that maximizes the likelihood of the speaker sequence
- The classification rate is 75%

From Classification to Diarization

	Baseline	Structure
Speaker Error	14.6%	12.7%

- Speaker diarization aims at finding in an audio stream “who is speaking when” and is based purely on acoustic information
- The detected structure can be included as a-priori information in the speaker diarization system
- Comparative improvements are obtained also on AMI meetings.

Future Work

- In the framework of the SSPNet, 5 meetings of the AMI corpus have been annotated in terms of social roles (following Bales and Slater): Protagonist, Gatekeeper, Supporter, Neutral
- These roles are better correlated with engagement in social interactions and will be investigated in the same vein as the others presented so far

Part III

- Networking

Liaison with Affective Sciences

- The activities in Track 1 involve both NCCRs:
 - IM2: A.Vinciarelli (G.Mohammadi)
 - Affective Sciences: M.Mortillaro
- Ongoing activities in Track 2 involve both NCCRs:
 - IM2: A.Vinciarelli
 - Affective Sciences: M.Mehu
- K.Scherer will be key-note speaker at an event On political speech co-organized by A.Vinciarelli (Rome, 10-12 Nov. 2010)

Liaison with SSPNet

- H. Bourlard (IM2) and A. Vinciarelli (IM2+SSPNet) co-chairs of the “*International Workshop on Mobile Social Signal Processing*” (Lisbon, Sep. 7th 2010)
- H. Bourlard (IM2) and A. Vinciarelli (IM2+SSPNet) co-editors of “*Mobile Social Signal Processing*”, LNCS, Springer (2011)
- L. Van Gool (IM2) involved in the “*Workshop on Socially Intelligent Surveillance and Monitoring*”, co-chaired by A. Vinciarelli (IM2+SSPNet)

Dissemination Events

- A. Vinciarelli (IM2+SSPNet)
 - Key note at “COST 2102 Training Event” - Caserta (Italy), Mar. 15-19, 2010
 - Key note at “Symposium on Unveiling Affective Signals” Eindhoven (NL), Aug. 27th 2010
 - Invited speaker at “Workshop on Society, Culture and Language” – Plymouth (UK), Nov. 10-12, 2010
 - Invited speaker at “Human Communication Dynamics” – Los Angeles (USA), Aug. 4-6 2010
 - Invited teacher at “Summer school on SSP” – Sestri Levante (Italy), Jul. 18-23 2010

Thank you for your attention!



FONDS NATIONAL SUISSE
SCHWEIZERISCHER NATIONALFONDS
FONDO NAZIONALE SVIZZERO
SWISS NATIONAL SCIENCE FOUNDATION

