Cognitive and Affective Influences of Voice-based Interactions in Immersive Experiences

PhD student: **Olga Chatzifoti**

Supervisor: **Dr. Maria Roussou**

Department of Informatics and Telecommunications (DI)

National and Kapodistrian University of Athens (NKUA)

Investigation area

How does **voice**

as an input modality

impact the **user experience**

in **immersive applications**?

Investigation area

How does voice

as an input modality

for **storytelling** and **roleplaying** purposes

impact the user experience

in terms of **critical** & **emotional engagement**

with the **text**

in immersive applications?

Research questions

How do

voice-based interaction methods (**VBIs**) compared to

silent interaction methods (controller-based)

affect:

Research questions

How do

voice-based interaction methods (**VBIs**)
compared to

silent interaction methods (controller-based)

affect:

- **Sense of embodiment?**

- **Cognitive response?**

- **Affective response?**

Research questions

How do

voice-based interaction methods (**VBIs**) compared to

silent interaction methods (controller-based)

affect:

- **Sense of embodiment?**
  - Sense of being, having & controlling a body in virtual reality
- **Cognitive response?**
- **Affective response?**

Research questions

How do

voice-based interaction methods (**VBIs**) compared to

silent interaction methods (controller-based)

affect:

- **Sense of embodiment?**

- **Cognitive response?**
  - Depth of an actor's investment interacting with a digital system.

- **Affective response?**

Research questions

How do

voice-based interaction methods (**VBIs**) compared to

silent interaction methods (controller-based)

affect:

- **Sense of embodiment?**

- **Cognitive response?**

- **Affective response?**

  - Intensity & direction of emotion (valence & arousal)

Hypothesis

VBIs can facilitate the user's **capacity to relate to the content**.

Read instead of listen to the author's diary?

By ARTE France

Hypothesis

For *some* user groups and with *some* design constraints,
VBIs can facilitate the user's capacity to relate to the content.

Read instead of listen to the author's diary?

By ARTE France

Expected outcomes

Provide **designers** with **recommendations** and **considerations** for integrating **VBIs** in **immersive experiences.**

-user profile groups

-synergetic design patterns

-opposing design patterns

-technological considerations

-physical setup

-future projections

Where can spoken thoughts go?

By Lucas Rizzotto

Significance

Promote VBIs in immersive experiences where relation to the text is crucial for their success, e.g. in application domains, like

arts & culture,

Interactive narratives

education

Actor practice

healthcare

Psychotherapy session

Image credit: Gemini advanced

Significance

Extend findings to social **human-human communication**?

RecRoom

Meta Horizon

Research methodology

Design a series of three cascading experiments with progressively complex VBIs: from voice to speech to conversation. (Fig. 1).
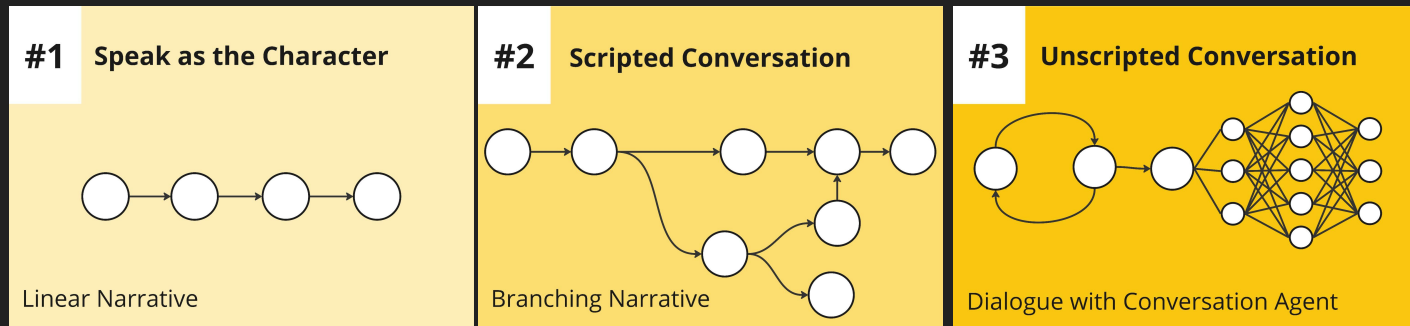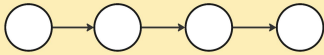

#1 Speak as the Character	#2 Scripted Conversation	#3 Unscripted Conversation
Linear Narrative	Branching Narrative	Dialogue with Conversation Agent

Figure 1. Series of planned experiments

Experiment series

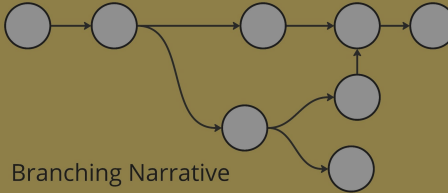## #1 **Linear Narrative**

Read out loud to unfold the narrative.

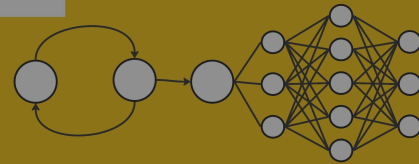Monologue excerpts sourced from (public access) books.

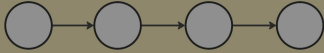| #1 Speak as the Character	#2 Scripted Conversation	#3 Unscripted Conversation
Linear Narrative	Branching Narrative	Dialogue with Conversation Agent

"I am …."

Experiment series

#2 **Branching Narrative**
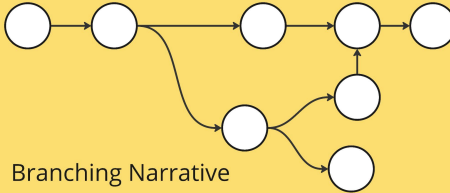
Read out loud to select an option in a dialogue with an NPC.

Commonly-used design pattern in role-playing.

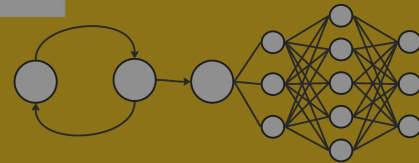| #1 **Speak as the Character**	#2 **Scripted Conversation**	#3 **Unscripted Conversation**
Linear Narrative	Branching Narrative	Dialogue with Conversation Agent

1. "X."
2. **"Y."**
3. "Z."
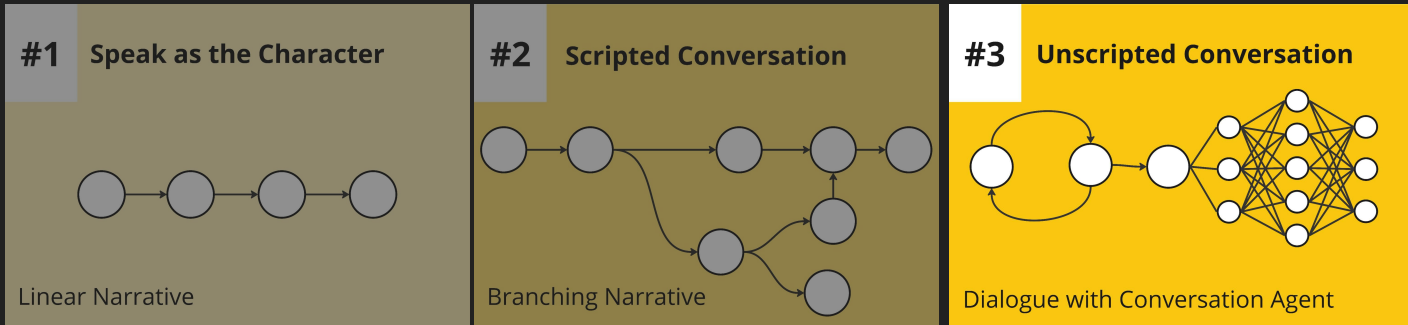
Experiment series

#3 **Emergent Narrative**

Speak with free-form natural language in a conversation with an AI-NPC.

Increasingly investigated design pattern for future adoption.

#1 Speak as the Character

Linear Narrative

#2 Scripted Conversation

Branching Narrative

#3 Unscripted Conversation

Dialogue with Conversation Agent

Control groups

Silent baseline (A) & 1 or 2 speaking test conditions (B/C)

Figure 2. Baseline (A) and test conditions (B,C) for Experiment #1

Differences

Self-reporting: account for personality differences

→ 1) User profiling

- ◆ **voice-related skills** (singing, performing, public speaking, etc)
- ◆ **preferences to content/literary genre** (sci-fi, fantasy, etc)
- ◆ **personality traits** (affective, empathetic, etc)

→ 2) Within-subject design

Differences, influences

Account for random parameters and biases

→ randomized condition order against **sequence bias**

→ randomized text per condition against **novelty bias** (within group)

→ standardized questionnaires on basic UX constructs, like SUS, Cybersickness, to account for **external influences**.

→ report related to **system performance** (accuracy, latency, robustness)

Image credit: Gemini advanced

Differences, influences and ambiguities

Account for unreliability of self-reporting

→ Add **biometric data** gathering

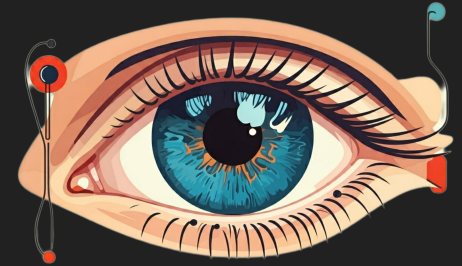→ **eye-tracking** → pupil dilation
(maybe fixations & saccades too?)
Perform user-based calibration.
Compare with eye-tracking data with self-reported data.
Disambiguate eye-tracking data.

→ **heart rate?**
→ **galvanic skin response?**

Voice says it all...?

Acoustic-prosodic features as a complementary way of estimating emotional/cognitive fluctuations

- Pitch (frequency)
    - Intonation (differences on pitch over time)
- Volume (amplitude)
    - Stress (differences on volume over time)
- Timbre (tonal quality or color)

Voice recording during experiment.

Offline Processing with Emotion Classification Models.

Multi-source data analysis.

Thank you!

Olga Chatzifoti, PhD candidate
olchatz@di.uoa.gr
Dr. Maria Roussou
mroussou@di.uoa.gr

HELLENIC REPUBLIC
National and Kapodistrian
University of Athens
EST. 1837