

Perceived Distraction and Annoyance in a VR Listening Task

Konstantin W. Kühlem¹, Aaron Welker¹, Agnes Widera¹, Sabine J. Schlittmeier², Janina Fels³, Andrea Bönsch¹

¹Visual Computing Institute, ²Work and Engineering Psychology, ³Hearing Technology and Acoustics, all RWTH Aachen University

Introduction

Background

- Embodied Conversational Agents (ECAs) in VR enable, e.g., education & collaboration, yet interaction succeeds only when users stay focused on the ECAs' speech
- Any competing stimulus acts as a *distractor* [Lavie2005]
 - ... impairing speech comprehension
 - ... distracting (attentional capture) or annoying (subjective irritation) users → dissociable constructs? [Massonnié2022]

Listening Task

- Open-plan office with two ECAs following a social-interaction model (gestures, turn-taking, periodic eye-contact) [Ehret2023]
- Heard-Text-Recall paradigm [Schlittmeier2023] in a dual-task fashion originally designed to assess memory performance
 - 1st task: ECAs narrate brief family stories [Erment2023] followed by nine content-related questions
 - 2nd task: keyword-detection while listening
 - left-controller press: family-related words
 - right-controller press: location-related words



Fig. 1: Office with two narrating ECAs amid multiple distractors

Overall Research Gap

- Lack of a unified taxonomy for distractors (in VR)
- Lack of systematic evaluation of how distinct distractor types affect user experience, task performance, and social presence

Research Objective

- Exploring how a subset of distractor types influences perceived distraction and annoyance during an ECA-driven listening task

Subset of Distractor Types

- Factors to be considered for a taxonomy
 - Intrinsic: properties describing the stimulus itself, e.g., sensory modality, intensity, social interaction, temporal pattern (repetition, duration, abruptness, ...)
 - Extrinsic: properties capturing how stimulus relates to scene, e.g., environment (originates inside VR or from real-world), congruence, user involvement, cognitive load, multimodality coupling
- Chosen subset with concrete stimuli for 1st exploratory study:

Congruence Social Int.	Office-Specific	Everyday	Extraordinary
Emitted by Virtual Agent (VA)	- Typing on keyboard - Chatting with co-worker	- Walking-by - Dropping glass	
Emitted by Object	- Screen notification - Printer printing - Phone ringing & blinking	- Ticking clock - Coffee grinder - Flying housefly - Blinking light	- Car crash - Fire alarm - Toy car

First Exploratory Study

Design

- Within-subject with six conditions
 - Baseline with no distraction,
 - ... followed by a randomized order of the five distractor types

- 5 participants (3M/2F, 31.4 ± 13.01 years)
- Per-condition subjective measurements:
 - 7-point Likert item for distraction
“How strongly did you feel distracted from the conversation by events in the office environment?” (1 - not at all, 7 - very much)
 - 11-point Likert item for annoyance (adapted from [Fields2001])
“To what extent did the background activity in the office disturb or bother you?” (0 - not at all, 10 - very much)

Cautious Findings

- Most distracting: Extraordinary stimuli
- Most annoying: Object-centered, office-specific stimuli

Cond.	Distraction M ± SD	Med.	Rank	Annoyance M ± SD	Med.	Rank
Baseline	1.00 ± 0.00	1	6	0.00 ± 0.00	0	6
VA / Office	1.80 ± 0.84	2	5	1.00 ± 1.00	1	5
VA / Everyday	2.60 ± 0.55	3	4	2.40 ± 1.14	2	4
Object / Office	4.20 ± 0.84	4	2	4.80 ± 1.79	5	1
Object / Everyday	2.80 ± 0.45	3	3	3.80 ± 0.84	4	3
Object / Extraordinary	4.80 ± 0.84	5	1	4.40 ± 0.55	4	2

- Annoying stimuli → verbal complaints
- Distracting stimuli → head-turns & flinching, indicating attentional capture

Take-Away

- Initial hints that distraction and annoyance diverge in VR contexts → basis for further hypotheses

Literature

Ehret et al., 2023
Who's next? Integrating Non-Verbal Turn-Taking Cues for Embodied Conversational Agents.
In *ACM Int. Conf. Intell. Virtual Agents*. doi:10.1145/3570945.3607312

Erment et al., 2023
AuVIST - An Audio-Visual Speech and Text Database for the Heard-Text-Recall Paradigm. doi: 10.18154/RWTH-2023-05543

Fields et al., 2001
Standardized General-Purpose Noise Reaction Questions for Community Noise Surveys: Research and Recommendations
In *Journal of Sound and Vibration*. doi:10.1006/jsvi.2000.3384

Lavie, 2005
Distracted and Confused?: Selective Attention Under Load. In *Trends in Cognitive Sciences*. doi:10.1016/j.tics.2004.12.004

Massonnié et al., 2022
Learning in Noisy Classrooms: Children's Reports of Annoyance and Distraction from Noise are Associated with Individual Differences in Mind-Wandering and Switching Skills. In *Environment and Behavior*. doi:10.1177/0013916520950277

Schlittmeier et al., 2023
Measuring Text Comprehension and Memory: A Comprehensive Database for Heard Text Recall (HTR) and Read Text Recall (RTR) Paradigms, with Optional Note-Taking and Graphical Displays. doi:10.18154/RWTH-2023-05285