Abstract ARSAD 2023

Spatiotemporal description of events in AD - the role of cognitive aspects and voice quality

Johansson, Roger, Holsanova, Jana, Lyberg Åhlander, Viveka, Sombeck Erika & Rastegar, Tina

To make visual media available to a non-sighted visually impaired audience, a sighted interpreter can provide audio description (AD). AD is a verbal description of visual events, where the spoken information intends to evoke "mental images" and to enhance meaning-making. The goal of AD is to increase the accessibility of visual information and to provide a non-sighted audience with a richer and more detailed understanding and enjoyment of, for instance, films and TV-shows.

The aim of the present study was to systematically investigate how the specificity of describing visuospatial and temporal state-of-affairs influences imageability and understanding for the non-sighted end users of AD, and to scrutinize how this interplay relates to the voice quality of the verbal narrator. For instance, in a visuospatial description, a woman could be described as either sitting "in front of" or "opposite" a man in a train compartment. The expression "in front of" contains less specified information about the visuospatial state-of-affairs than "opposite", as it does not reveal whether the woman has her back towards the man or whether she is facing him. Similarly, descriptions of temporal change contain varying degrees of specificity through the manner of motion. For example, a boy can be described as either "entering" the classroom door or as "rushing through" the classroom door. The expression "rushing through" here carries more specified information about the described state-of-affairs.

In our experiment, a group of 21 non-sighted participants (congenitally blind or who lost their sight early in life) and a group of 21 sighted control participants listened to 55 short scenarios describing temporal change or visuospatial relations with varying degrees of specificity. The two groups were matched in respect to verbal working memory capabilities (through the Competing Language Processing Task - CLPT). After each scenario the participants were to evaluate on a scale 1-6 (1) how well they could imagine the content of the scenario; (2) their overall comprehension of the scenario; (3) how effortful it was to listen to the scenario; and (4) how pleasant it was to listen to the scenario. Across the whole experiment the participants listened to an equal number of scenarios narrated in a typical voice and in a dysphonic voice (counterbalanced across specificity conditions).

Results revealed that more specified scenarios of temporal change increased imageability and comprehension for the non-sighted participants, but not for the sighted control group. For the visuospatial scenarios, the overall pattern was less clear. But there was a tendency for more specified scenarios to increase imageability and comprehension, independent of sightedness. Additionally, scenarios narrated with the dysphonic voice were found to increase listening effort and to decrease pleasantness for the non-sighted group, but not for the sighted control group.

Collectively, these results demonstrate the importance of considering fundamental cognitive and psycholinguistic aspects of spatiotemporal description of events for successful AD. They also exhibit an intricate relation to voice quality and how it affects perception, understanding and enjoyment.