Audio, visual or audio-visual text processing in online education

Tatiana Petrova
Institute for Cognitive Studies, St Petersburg State University, Russia Federation

https://doi.org/10.36505/ExLing-2022/13/0038/000580

Abstract

The purpose of the study was to check if there are any differences in processing the same block of information presented in different text formats. In Exp.1 we tried to reveal the influence of leading perception channel of a person (audio channel or visual channel) on the effectiveness of processing written and oral texts. Exp. 2 was aimed to study the peculiarities of comprehension and memorising different types of the texts: (1) written text-only, (2) audio-only, (3) written text + infographics, (4) audiotext + infographics, (5) written text + infographics + audiotext. As a result, a hierarchy of information presentation formats was built in terms of the success of their perception and understanding.

Keywords: text processing, text format, audiotext, verbal text, multimodal text, Russian

Introduction

The text type is among the readability categories and it influences the effect of reading perspective. The increased interest in online education causes an issue to find out the most efficient forms to present the information. Developing the best compromise and form of text presentation in oral and written form becomes especially relevant in connection with the move to remote learning, in which many formats of interaction between the teacher and students involve a combination of auditory and visual modalities.

Primary focus of the present study is on the fundamental scientific problem of verbal and non-verbal information perception, conjugated with methods for studying human perception and understanding of text content (Mayer, 2009). In particular, of essential interest is how the text format is related to the text comprehension quality. Paivio in his works points out that information is better stored in memory if presented by text and the corresponding illustration, rather than by text only (Paivio, 2006). On the other hand, the accompanying verbalization of the material itself presented and perceived in the form of images can provoke its distortion and/or forgetting during subsequent recall of particularly figurative information.

According to the Cognitive Load Theory (Sweller, 2003) our cognitive architecture includes sensory, working, and long-term memory and there are separate systems to process verbal and non-verbal information; comprehension occurs when a recipient selects relevant information from each memory and

organizes it into a coherent representation and makes appropriate connections between the presentations of each memory. Thus, the more independent the elements of information (text, sound, image), less cognitive work will be done by the person to choose which modality (verbal or visual) he prefers in processing this text.

There are three main factors that influence on the process of understanding the educational material (Sweller, 2003): 1) effect of split attention (it is easier to a person not to share his attention between different sources of the same modality: for example, verbal text and picture); 2) the modality effect: the text (oral or written) cannot just recapitulate the image, or visa versa; there is a partial autonomy of the visual and auditory modalities); 3) the redundancy effect (redundancy of mental/physical activity, redundancy of detail/abstract, visual/auditory redundancy). Moreno and Mayer (2000) added three main principles to the cognitive load theory: principal of spatial contiguity, principal of temporal contiguity, principle of coherence.

Text comprehension also depends on a number of individual characteristics of a reader: his language skills, reading experience, working memory, background knowledge. Leading perceptual modality of a recipient seems to be an important factor, that influences on the process of text understanding. There are several methods to determine perception channel of a person. For Russian native speakers it is a questioner proposed by Efremtseva (2018). Participants have to answer 48 questions revealing how it is easier for them to perceive information, thus to determine their leading channel: auditory, visual or kinesthetic.

Goals

This study aims to answer two questions: 1) are there any differences in processing written, oral and multimodal texts by Russian people with different perception channels? 2) what type of a text format contributes to more successful information processing and understanding.

Experiment 1

In Exp. 1 we investigated how 52 native speakers of Russian (20 males and 32 females), aged 18–25 y.o., process, understand and remember information when online reading and online listening oral and written texts. First, we checked the preferred perception channel of the participants, using the special Russian questionnaire designed by S. Efremtceva (https://onlinetestpad.com/ru/test/1361-diagnostika-dominiruyushhej-perceptivnoj-modalnosti-s-efremceva). As a result, we chose 15 participants with visual perception channel and 15 participants with audio perception channel. These 30 students took part in the main experiment. Educational texts for Russian students from http://www.orator.ru/ were used as the material. All

the texts were of the same length and the same level of readability (checked via http://readability.io/).

In a two-group experimental design, the participants examined four different texts in two different formats online. Afterwards, they answered the factual and analytical questions, and estimated the difficulty of each text. The experiment was conducted online. The results obtained by question-answer technique show no significance of the text type factor ($\beta=-1.72$, SE=0.29, p=2.91e-09), perception channel factor ($\beta=-1.55$, SE=0.27, p=1.71e-08), and of the interaction between these two factors ($\beta=-1.58$, SE=0.33, p=3.12e-06). Though we do revealed the tendency for visuals to estimate oral texts as more difficult in comparison with written texts. On the contrary, audials (participants with audial perception channel) gave higher marks of understanding audio texts.

Experiment 2

In Exp. 2 we hypothesised that multimodal text is the most efficient and easiest for comprehension text format. In a five-group experimental design, 50 Russian participants (12 female, aged 18-30 y.o.) examined five different texts from PISA (https://www.oecd.org/pisa/) in five different formats: (1) written text-only, (2) audio-only, (3) written text + infographics, (4) audiotext + infographics, (5) written text + infographics + audiotext. It was the same block of information in each of the formats. The statistical analysis was done in the R programming environment. To analyze the subjective scaling data, we used ordinal logistic regression. We revealed significant differences between all text formats except the pair (3) – (5).

Linear Hypotheses:

	Estimate	Std. Error	z value	Pr(>|z|)
2 - 1 == 0	1.2422	0.2334	5.321	6.18e-07 ***
3 - 1 == 0	2.5506	0.2762	9.235	< 2e-16 ***
4 - 1 == 0	0.5372	0.2257	2.380	0.03460 *
5 - 1 == 0	2.2787	0.2632	8.657	< 2e-16 ***
3 - 2 == 0	1.3084	0.2732	4.789	8.39e-06 ***
4 - 2 == 0	-0.7050	0.2310	-3.052	0.00681 **
5 - 2 == 0	1.0365	0.2601	3.985	0.00027 ***
4 - 3 == 0	-2.0133	0.2725	-7.387	1.20e-12 ***
5 - 3 == 0	-0.2719	0.2918	-0.932	0.35156
5 - 4 == 0	1.7415	0.2596	6.709	1.37e-10 ***

Audio-only text was the most difficult format to comprehend (42% right answers to after-the-text questions). The format 'written text + infographics'

and the format 'written text + infographics + audiotext' were the easiest to comprehend (86% and 83% right answers respectively).

Conclusion

The overall results made it possible to build a hierarchy of text presentation formats that are the most successful for recipients in online education (from the most effective to the most difficult): 'written text + infographics' / 'written text + infographics + audiotext' – 'written text' – 'audio text + infographics' – 'audio-only'.

Our findings confirm the theory of D. Sweller (2003), according to which the greater the cognitive load experienced by the recipient when studying the text, the more successfully he understands this text. In our study, the success of understanding means the correct answers to questions after the text. Our results show that the so-called multimedia effect helps to integrate the new information in the cognitive system and to remember, though it is a difficult process, especially for non-native speakers. The obtained results also confirm the theory of R. Mayer and R. Moreno [2000] about the parallel presentation of information in several codes, which increases the success of information processing.

The results of the work can be taken into account during making the materials for online educational resources.

Acknowledgements

The study was supported by the research grant no. 21-18-00429 "Cognitive mechanisms of multimodal information processing: text type & type of recipient" from Russian Science Foundation. I am especially grateful to my student Eugenia Smelova for data collection and to Varvara Magomedova for her help with statistical analysis.

References

Mayer, R.E. 2009. Multimedia Learning, 2nd Edn Cambridge, MA: Cambridge University Press.

Moreno, R., Mayer, R. E. 2000. A learner-centered approach to multimedia explanations: Deriving instructional design principles from cognitive theory. Interactive multimedia electronic. Journal of Computer-Enhanced Learning. http://imej.wfu.edu/articles/2000/2/05/index.asp

Paivio, A. 2006. Dual Coding Theory and Education. Draft Chapter for the Conference on Pathways to Literacy Achievement for High Poverty Children. Ann Arbor, MI: The University of Michigan School of Education.

R Core Team 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Sweller, J. 2003. Cognitive Load Theory: A Special Issue of Educational Psychologist. LED, Inc.