学 位 論 文 内 容 の 要 旨

博士の専攻分野の名称　博士（情報科学）　氏名　張　慧丹

学 位 論 文 題 名

Visual Factors Shaping Text Legibility and Preference in Virtual Reality: A Comparative Study of Native and Non-Native Speakers

（バーチャルリアリティー環境におけるテキスト判別性と嗜好性に影響する視覚要因の調査: 母語話者と非母語話者の比較研究）

Everyday interaction is largely driven by fast, unconscious perception rather than deliberate reasoning. For text, this means that legibility must be supported at a perceptual level before higher-level comprehension can unfold. Classical work on print and 2D displays has demonstrated how type size, line length, contrast, and typeface style impact reading performance; however, these findings were obtained under specific viewing distances, display resolutions, and environmental conditions. Immersive near-eye displays (e.g., Head-mounted Display; HMD) in Virtual Reality (VR) change these physical constraints: HMDs have limited and non-uniform fields of view, lens distortion, dynamic blur from head motion, and spatially varying pixels-per-degree. As a result, parameters that were once considered “safe defaults” on paper or desktop screens, such as minimum text size, stroke width, or preferred font style, must be re-evaluated for VR. Many recent VR/AR studies have attempted to tune readability by manipulating text size, background, contrast, motion, or walking conditions; however, they employ different hardware, scripts, languages, and participant groups, resulting in findings that form islands of evidence rather than a unified, comparable basis for multilingual VR interfaces. At the same time, psycholinguistics and eye-movement research indicate that what constitutes “clear enough to read” is not a single threshold, but rather depends on both the reader and the script. First-language (L1) readers can lean more on top-down predictions and well-established lexical representations, which helps them compensate for imperfect visual input, whereas second-language (L2) readers depend more on bottom-up visual detail. Cross-script studies further indicate that logographic and alphabetic readers do not rely on the same visual cues when recognizing written forms. These findings (reviewed in Chapter 1) suggest that design rules for VR text should take language background and script into account, yet most VR legibility work still focuses on English or treats readers as a homogeneous group, and systematic L1-L2 comparisons in VR remain rare.

To address these gaps, this dissertation adopts a controlled, comparative strategy. All experiments are conducted on the same VR device, under the same laboratory environment, with carefully calibrated distance and text size, These visual variables are defined and systematically categorized in Chapter 2, and then consistently applied across all experiments. Within this common framework, this dissertation recruited native Japanese, English, and Chinese readers and asked them to perform the tasks in both their first language (L1) and a less familiar second language (L2), so that any observed differences could be attributed to visual text factors and language background rather than to hardware or contextual factors. Structurally, the dissertation follows a Multi-Trait Multi-Method matrix (MTMM) approach, focusing on two traits, legibility and preference, and measuring them with three complementary methods: a preferred viewing distance task in Chapter 3, a glyph identification task in Chapter 4, and a subjective font rating task in Chapter 5. This MTMM approach enables the identification of effects that are robust

across traits and methods versus those that are trait- or method-specific.

The preferred viewing distance task (Chapter 3) uses viewing distance as a first, intuitive measure of legibility in VR. Participants adjust the text panel in the headset closer or farther until it feels "clear and comfortable" to read. Because angular size changes with distance, the task captures a self-chosen combination of size and distance rather than a fixed threshold imposed by the experimenter. Because there is no single correct answer, the outcome is intentionally subjective and reflects each reader's internal criterion for legibility under identical device and environmental conditions. Overall, this method demonstrates that the preferred viewing distance in VR is shaped not only by device limitations but also by script-specific top–down compensation and differences in effective perceptual span between native and non-native readers.

The glyph identification task (Chapter 4) offers a more constrained, performance-based view by focusing on glyph identification. The experiment manipulates glyph complexity and similarity and asks whether readers can correctly match a briefly presented target to one of two options. Errors can thus occur even when the stimulus is visible, because highly similar alternatives invite confusions. The task yields objective accuracy and response-time measures. Overall, this method demonstrates that VR preserves deep cross-script differences in visual strategy: character-script natives(e.g., Chinese and Japanese) are most vulnerable to similarity-driven confusions, alphabetic-script natives are more constrained by complexity.

The subjective font rating task (Chapter 5) again takes a subjective perspective, evaluating legibility through font judgments. Participants compare short text samples presented at the same position, size, and content in VR but rendered in different fonts, and indicate which font feels more comfortable and easier to read. Brief informal interviews probe the reasons behind their choices. By holding distance and content fixed, this study isolates the role of font design in shaping subjective legibility and shows how readers experience different font options under realistic viewing conditions. Overall, this method reveals that subjective font choices in VR are jointly driven by legibility and aesthetics: native readers' comfort closely tracks configurations that support strong performance, whereas non-native readers more often favor familiar or visually appealing styles even when they are not objectively optimal.

The final discussion chapter (Chapter 6) discusses the three methods within the MTMM framework, treating them(Chapter 3-5) as complementary components of a unified comparative framework. Convergences across distance adjustments, identification performance, and subjective font ratings highlight general effects of script and language background on VR legibility, while discrepancies between performance metrics and preferences reveal trait-specific aspects of user experience. Building on this synthesis, the dissertation introduces a prototype Visual Text Legibility System (VTLS) implemented in Unity that encodes the convergent findings as simple rules and predictive functions for real-time, legibility-aware text rendering. Taken together, these three components of this work (the unified cross-script dataset, the MTMM-based comparison of legibility and preference, and the VTLS prototype) show that a common VR setup can support systematic comparisons across scripts, reader groups, and visual factors. They provide an initial foundation for multilingual, adaptive VR interfaces that respect users' perceptual limits and linguistic diversity.