

LyricBatheにおけるMR環境での歌詞体験に対するインタラクションプリミティブの影響の分析

小島 颯英^{1,a)} 中小路 久美代^{1,b)} 加藤 淳^{2,c)} 後藤 真孝^{2,d)}

概要：

LyricBathe は、HoloLens 2 を用いて楽曲の進行に同期した歌詞テキストを物理空間に重畳表示し、ユーザや身の回りの物理環境と影響する形での歌詞の体験を可能にする MR 環境である。これまでに、文字の大きさや身体に対する表示位置、文字の操作手法などのバリエーションを多様に組み合わせることで、計 24 個のモードを実装し、各モードにおいてユーザの注視箇所や頭部位置がどのように変化するかを定量的に観察してきた。本稿では、モード単位の評価ではなく、それぞれ複数個のモードに共通して現れているインタラクションデザインのバリエーション（「インタラクションプリミティブ」と呼ぶ）について、これらをユーザの注視行動から定量的に分析した結果を報告する。空間内での歌詞フレーズの表示位置、歌詞を構成する文字の大きさ、文字の表示継続時間、文字操作ジェスチャ、という 4 種のインタラクションプリミティブについての分析から、例えば、文字の大きさの変化によって、音楽とともに歌詞を読む体験から文字探索的な体験へと変化するという傾向が認められるなど、特徴的な体験傾向が示唆された。これらの知見は、MR 環境における歌詞体験のインタラクションデザイン指針の構築に資するものである。

1. はじめに

音楽に付随する歌詞を構成する文字表現は、楽曲の楽しみを視覚的に拡張し得る主要素の一つである。これまで、歌詞カードや歌詞をアニメーション表示するリリックビデオなど、静的あるいは映像的な文字表現が主流であった。近年では、加藤ら [1] による Lyric App Framework が提案され、楽曲の進行に同期して多様な歌詞体験を実現する「リリックアプリ」の開発が可能となっている。これら既存のアプローチの多くは矩形ディスプレイ上での視覚体験を前提としており、身体や物理環境との関わりを伴う歌詞体験の事例の報告はあまりみられない。

本プロジェクトで構築してきた LyricBathe は、光学透過式ヘッドマウントディスプレイ (HoloLens 2) を用いた MR 環境において、歌詞を物理空間内に 3DCG オブジェクトとして重畳表示し、ユーザの身体や周囲の物理的環境と相互作用することを可能にするシステムである (図 1)。ユーザは、手の動きや空間的な移動を通じて身体的に歌詞と関わりながら音楽を体験することができる。

キネティックタイポグラフィ [2], [3] など、動的テキスト表現に関する研究は 30 年以上行われているが、いずれも主に二次元空間での視覚的演出に焦点を当てていた。近年は、文字体験を MR (Mixed Reality) 環境へ拡張する研究が現れつつあり、たとえば横濱ら [4] による物理的な本の白紙ページの上に仮想の詩の文字を重畳表示する MAVenRead が挙げられるが、このような研究は、動的な音楽進行に同期して体験される歌詞表現とは性質が異なる。音楽体験の身体的拡張に関しては、Miyagawa ら [5] や Jeon ら [6], [7] による物理空間に音源を配置・制御する研究が存在する。これに対して LyricBathe は、音そのものではなく「歌詞」という意味的・言語的対象を身体的に扱う探究である。

LyricBathe プロジェクトではこれまでに、文字の大きさや表示位置などを変えながら、さまざまな歌詞表示のパターンを組み合わせた計 24 個のモードを、キーボードを介して変更可能なツールとして実装し [8]、それぞれのモード体験時のユーザの注視位置計測と頭部位置計測データを用いて、誘発される注視行動や身体的挙動の違いを分析した [9]。

これに対して本稿では、モード単位の評価ではなく、それぞれ複数個のモードに共通して現れているインタラクションデザイン（表現系と操作系を構成する要素）のバリエーション（インタラクションプリミティブ; 2.2 参照）に

¹ 公立はこだて未来大学

² 産業技術総合研究所

a) g2124009@fun.ac.jp

b) kumiyo@fun.ac.jp

c) jun.kato@aist.go.jp

d) m.goto@aist.go.jp

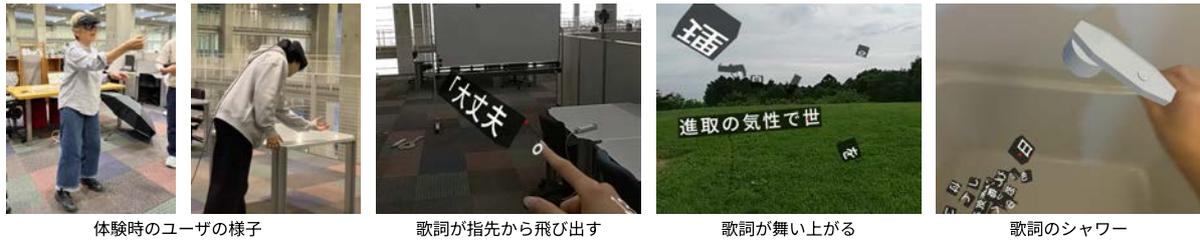


図 1 LyricBathe での歌詞へのインタラクション

ついて、ユーザの注視行動から定量的に分析した。空間内での歌詞フレーズの表示位置、歌詞を構成する文字の大きさ、文字の表示継続時間、文字操作ジェスチャ、という4種のインタラクションプリミティブについて、それらが注視割合の構成に与える影響、視線遷移構造に現れる体験様式の差異をユーザ観察データから明らかにした。MR環境における身体的歌詞体験を構成する要素レベルのデザイン指針の導出に寄与するものと考えている。

2. LyricBathe とインタラクションプリミティブ

本章では、LyricBathe の構成と、本研究で分析の対象とするインタラクションプリミティブを説明する。

2.1 LyricBathe の概要

LyricBathe は、光学透過式ヘッドマウントディスプレイ (HoloLens 2) を用いて、楽曲の進行に合わせて歌詞テキストを三次元的に空間へ配置し、身体動作を通じて歌詞と関わる体験を可能にする MR システムである。ユーザは、空間に浮かぶ歌詞を眺めたり、身体や手の動きを通して歌詞に触れたりすることに加えて、文字をつまんで部屋の写真立ての前に積み上げたりするなど、音楽と歌詞を一体的に体験することができる (図 1)。本システムでは、音楽の音源データ (MP3 形式) と対応する歌詞データ (リリックビデオ制作支援サービス「TextAlive」[3] により生成された発声タイミング情報が各文字に付与された CSV) を読み込み、文字オブジェクトの表示を制御する。これにより、音楽の時間構造と連動した視覚的・身体的体験を可能としている。

これまで、歌詞を構成する文字サイズの大小、表示継続時間の長短、異なる表示位置やインタラクション手法など、それぞれのバリエーションを多様に組み合わせて、「文字が指先から飛び出し漂い続ける」「現れた巨大な文字が地面に積み上がる」といった、合計 24 個のモード (M01～M24) を構築している [9]。

2.2 LyricBathe におけるインタラクションプリミティブ

前述のように本稿では、それぞれいくつかのモードに共通して現れている (使用されている) インタラクションデ

ザインの要素 (インタラクションプリミティブ) を定量的に評価することを目指した。ここでインタラクションプリミティブとは、TbVP (Time-based Visual Presentation) Landscape [10] の考え方を発展させたものである。TbVP は、映像体験を構成するツールのインタラクティブ性の構成要素を時間的特性 (temporal 軸) と視覚的特性 (visual 軸) に分類し、インタラクティブな映像体験アプリケーションをそれらの組み合わせで構成するものである。これに対して LyricBathe の体験空間は、TbVP が提示した視覚的特性を空間的構成 (spatial 軸) として捉え、時間軸に加えて身体的相互作用 (身体性の次元: embodiment 軸) が加わった三次元空間として考えることが出来る (図 2)。LyricBathe のインタラクションプリミティブは、MR 環境における「時間・空間・身体」の関係进行操作するための構成的なインタラクションデザイン要素となる。この意味において LyricBathe の 24 個のモードは、単に視覚効果の違いを示すに留まらず、ユーザの身体的関与の仕方を異ならせる構成的要素 (インタラクションプリミティブ) の組み合わせとしてデザインされている。

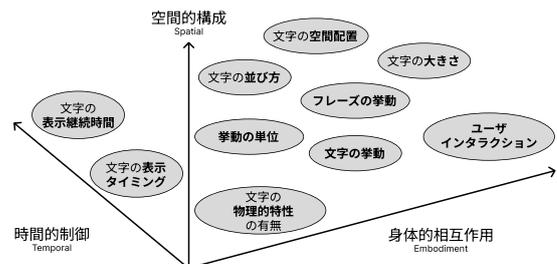


図 2 インタラクションプリミティブの三軸モデル

今回分析の対象としたのは、以下に示す 4 種類のインタラクションプリミティブである (表 1)。それぞれにバリエーションがあり、実装した 24 個のモードのうち、複数個のモードに共通して現れるものである。各モードは、これらの組み合わせで実現されている。

(1) *display positioning*: 空間内での歌詞フレーズの表示位置を規定する。空間内の定まった位置に固定表示 (absolute)、ユーザの指の位置やオブジェクトに追従表示 (relative)、空間配置の一軸のみをオブジェクトなどの動きに追従 (absolute と relative の両方) とい

表 1 LyricBathe におけるインタラクションプリミティブとそれに関わるモード

Interaction Primitives		バリエーション	関連するモード
1	display positioning	空間内での歌詞フレーズの表示位置を規定 ・空間内の定まった位置に固定表示 (absolute) ・ユーザの指の位置やオブジェクトに追従表示 (relative) ・空間配置の一軸のみをオブジェクトなどの動きに追従 (absoluteとrelativeの両方)	・ absolute: M03, M07, M09, M10, M14, M15 ・ relative: M01, M08, ・ absoluteとrelativeの両方: M23, M24
2	scaling character	歌詞を構成する文字の大きさを規定 ・基準となる大きさ (1倍) ・5倍の大きさ ・15倍の大きさ ・文字ごとにランダムなサイズ (random scaling)	・1倍: M01, M03, M07, M08, M14, M15, M23, M24 ・5倍: M09 ・15倍: M10
3	display duration	文字の表示継続時間を規定 ・フレーズの歌唱終了と同時に表示終了 (disappeared at the end of each phrase) ・フレーズの歌唱終了後、25秒間表示 (25 seconds after the end of the phrase)	・フレーズと同時に表示終了: M01 ・25秒表示: M03, M07, M08, M09, M10, M23, M24
4	user interaction	表示された歌詞オブジェクトなどに対してユーザが行うことが出来るインタラクションを規定 ・指先の位置に追従したインタラクション (finger-based gestures) ・手のひらでの接触やオブジェクトの把持によるインタラクション (palm-based gestures) ・シャワーヘッドなどのインタフェースオブジェクトによる文字の操作 (tool-based)	・finger-based: M01, M23, M24 ・palm-based: M03, M07, M09, M10, M14, M15, M23, M24 ・tool-based: M08, M15

たバリエーションを有する。

- (2) *scaling character*: 歌詞を構成する文字の大きさを規定する。基準となる大きさ (1倍), 5倍の大きさ, 15倍の大きさ, 文字ごとにランダムなサイズ (random scaling) といったバリエーションが可能である。
- (3) *display duration*: 文字の表示継続時間を規定する。フレーズの歌唱終了と同時に表示終了, フレーズの歌唱終了後に 25 秒間表示といったバリエーションを有する。
- (4) *user interaction*: ユーザが表示された歌詞オブジェクトなどに対して行うことが出来るインタラクションを規定する。指先の位置に追従したインタラクション (finger-based gestures), 手のひらでの接触やオブジェクトの把持によるインタラクション (palm-based gestures), シャワーヘッドなどのインタフェースオブジェクトによる文字の操作 (tool-based) といったバリエーションがある。

なお, LyricBathe には, これら 4 種以外にも, 文字の並び方や表示のタイミングなどのインタラクションプリミティブがあるが, それらの分析は今後の課題となる。

3. ユーザ体験データの分析手法

LyricBathe プロジェクトでは, これまでに実施したユーザ観察実験 [9] を通して, 参加者・モード・楽曲の 3 変数による体験の違いを分析し, モードによって注視行動が一貫して誘発される場合と, 個人差が表れやすい場合があることを示した。

本稿では, 同じ分析の枠組みを維持しつつ, インタラクションプリミティブという構成要素の観点から, ユーザ体験データを分析する。分析には, これまでに収集した以下の二種類の実験データを用いた。それぞれの条件の組み合わせごとに, 注視箇所の時系列データが記録されている。なお, 実験データの一部は, [9] に詳述したユーザ体験デー

タと同一のものであるが, 分析の観点が異なる。

- 1 名の参加者による全 24 モード実験データ
 - 1 名の実験参加者
 - 全 24 モード (M01–M24)
 - 1 楽曲 (1 分 18 秒, Suno [11] により生成)
- 複数参加者 (4 名) による 8 モード実験データ
 - 4 名の実験参加者 (P01, P02, P03, P04)
 - 8 モード (M01, M03, M07, M08, M09, M10, M14, M15) (図 3)
 - 同一歌詞を用いて Suno [11] により生成した曲調の異なる 3 楽曲 (47 秒, ロック, ヒップホップ, バラード)



図 3 分析対象データ取得に用いた 10 種のモード

本分析では, 着目する 4 種類のインタラクションプリミ

タイプに関わる 10 個 (M01, M03, M07, M08, M09, M10, M14, M15, M23, M24) のモード (図 3) の注視データをもとに、各モードに含まれるインタラクションプリミティブごとにそれらと比較することでその影響を検討した。なお、注視箇所の時系列データでは、以前の分析 [9] と同様に、注視対象を以下の 6 つのタグで分類し注視時間を記録している。

- CCO (Current Character Object): その時点で発声されている歌詞の文字上または表示から 1 秒以内の歌詞オブジェクト上に視線がある
- CPO (Current Phrase Object): その時点で発声されている歌詞の文字が入ったフレーズ上に視線がある (CCO 以外)
- TLO (Touched Lyric Object): 手指で操作している歌詞の上に視線がある
- OLO (Other Lyric Object): それ以外の歌詞の文字の上に視線がある (CCO でも CPO でも TLO でもないもの)
- IVO (Interaction Virtual Object): 表示されているユーザインタフェースオブジェクト (シャワーキューブや矢印オブジェクト) の上に視線がある
- UNO (Undefined Observation): どの状態にも当てはまらない場所を注視している

4. 分析結果: インタラクションプリミティブが注視行動に与える影響

本章では、着目する 4 種類のインタラクションプリミティブがどのようなユーザのインタラクション体験を生じさせているかを考察するために、各注視タグの注視割合を分析した結果を論じる。本分析の焦点は、モード自体の主観の評価ではなく、モードに備わっているインタラクションプリミティブの差異がユーザの行動 (注視箇所のタグ別割合、視線遷移など) にどのような一貫した傾向や変化を生み出すかにある。複数のプリミティブが同時に関与する場合に、どのプリミティブがより強い影響を有するのかについても検討する。

4.1 文字の大きさによる注視傾向の変化

scaling character というインタラクションプリミティブの効果を分析することを目的として、それ以外は共通のインタラクションプリミティブを有する、文字サイズが 1 倍、5 倍、15 倍と異なる 3 つのモード (M03, M09, M10 図 3 参照) のユーザ体験データを比較した。その結果、文字サイズの違いが注視割合に与える影響が明らかとなった。

図 4 に、参加者 (P01 から P04) と、3 楽曲 (ロック調, ヒップホップ調, バラード調) で平均した注視箇所の時間的割合を示す。文字サイズが 1 倍 (M03) では、注視が CCO や CPO といった、発声中の歌詞やフレーズへ比較的

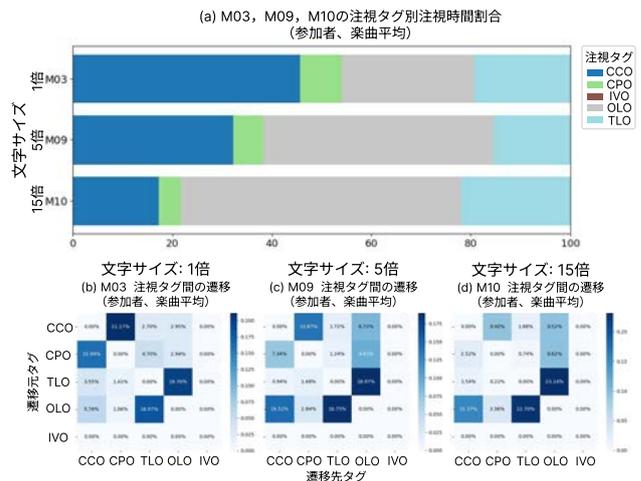


図 4 M03, M09, M10 の注視タグごとの注視行動の構成 (参加者 (P01 から P04), 楽曲 (ロック調, ヒップホップ調, バラード調) 平均)

集中しており、ユーザは音楽の進行に合わせて歌詞を「読む」体験をしている様子が見られた。一方、文字サイズが 5 倍 (M09) では、CCO 及び CPO の割合が減少し、TLO や OLO といった「落ちている歌詞」や「触れている歌詞」への注視割合が増加した。文字サイズが 15 倍 (M10) ではさらにこの傾向が強まり、OLO が大きな割合を占めるようになった (図 4)。また、注視タグ間の注視遷移に関するヒートマップを見ても、M03 と M09, M10 を比較すると文字の大きさが大きくなるほど CCO と CPO 間の遷移割合が減少し、OLO と TLO 間の遷移割合が増大していることがわかる (図 4)。

要約すると、文字の大きさを拡大することで、ユーザは音楽に同期して歌詞を追う行動から、周囲に存在する歌詞オブジェクトを探索したり触れたりする行動へとシフトする、ということになる。*scaling character* プリミティブは、注視割合を大きく変化させる強い要因であると考えられる。

4.2 文字表示位置による体験の違い

display positioning というインタラクションプリミティブに関しては、文字の出現位置を固定する *absolute positioning* と、ユーザの身体や手の動きに追従して動く *relative positioning* とでは、注視割合の構成に違いがあることが、それぞれのプリミティブを構成要素として擁するモードごとの観察データから明らかとなった。

relative positioning を含むモード (M01, M08 図 3 参照。以下同様) では、全参加者で似た注視割合の構成が得られた。これは、インタラクション体験が統制され、全体として均質な体験が生じることを示している (図 5)。

これに対して、*absolute positioning* を含むモード (M03, M07 など) では、参加者ごとに注視割合の構成が異なった (図 5)。これは、歌詞が空間内の特定位置に固定されてい

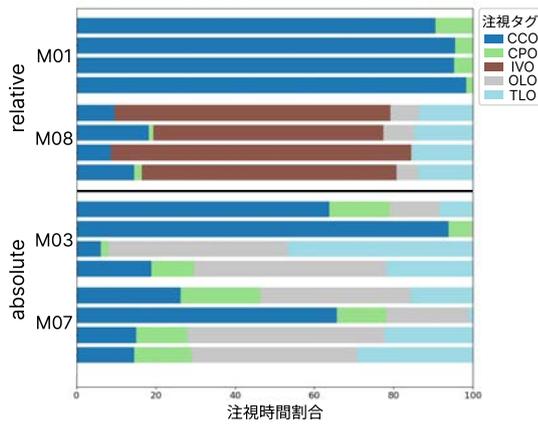


図 5 relative positioning を含むモード (M01, M08) と absolute positioning を含むモード (M03, M07) の参加者別 (P01~P04), 楽曲 (ロック調, ヒップホップ調, バラード調) 平均の注視割合の構成

るため, ユーザがそれぞれ異なる動き方で文字を追う必要があり, 探索的で個人差の大きい体験を生じさせていると考えられる。

これらの結果から, absolute positioning は相対的にユーザ依存のばらつきを生む強い要素であり, relative positioning は逆に統制された体験をもたらす方向に作用していることが示唆される。

4.3 文字表示継続時間の影響

display duration というインタラクションプリミティブについても, 注視割合に影響を与えていると考えられた。



図 6 表示継続時間による注視遷移傾向の変化。(b) は図 4 (b) の再掲

フレーズの終了とともに文字が消える短い表示時間のモード (M01) では, それ以降の歌詞オブジェクトの状態 (OLO など) に注視点が遷移しないため, 注視が CCO と CPO に集中し, 音楽に同期した注視行動がみられた。一方, フレーズ終了後も 25 秒間表示を継続するモード (M03) では, CCO や CPO 以外の注視タグへの注視割合が増加し, 音楽の余韻の中で歌詞を探索するような行動が観察された (図 6)。

このことから, 表示継続時間が長くなるほど, 「音楽とともに歌詞を読む体験」から「現在の歌唱から時間的に離れ

た歌詞の探索」を取り入れた体験へ変化することが示唆された。

4.4 操作タイプの影響と他プリミティブとの関係

操作タイプを表すインタラクションプリミティブ *user interaction* に関しては, 指先で直接文字とのインタラクションを行う *finger-based gestures* 型のモード (M01, M23, M24) と, 仮想オブジェクト (シャワーヘッドなど) を介して文字へのインタラクションを行う *tool-based* 型のモード (M08) を比較することでその違いが明らかとなった。

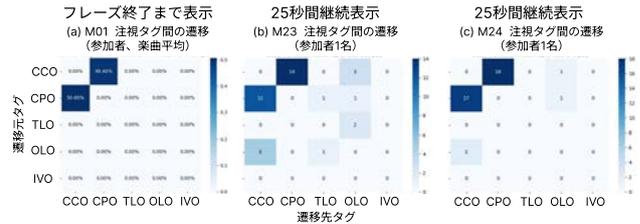


図 7 finger-based gestures 型のモードの遷移特徴 (遷移分布が CCO-CPO 間に集中)。(a) は図 6 (a) の再掲

finger-based gestures 型では, ユーザは文字へのインタラクションインタフェースである自分の指先をほとんど注視せず, 指先から飛び出した歌詞そのものを追っていた。ユーザは「指を見る」のではなく「文字を見る」ということがユーザの映像からも確認された。データからも, 表示時間が長くても短くても注視割合は CCO や CPO 中心に安定していることが示されていた (図 7)。

一方, tool-based 型の M08 では, 歌詞が出てくる仮想オブジェクト (シャワーヘッド) そのものを頻繁に注視し, IVO (操作オブジェクト) タグの割合が増加した。これは, 物理的な道具が手の延長として一体化する現実世界とは異なり, デジタルな道具の場合は「それを操作できているか」を確かめる注視行動が生じるためと考えられる (図 8)。

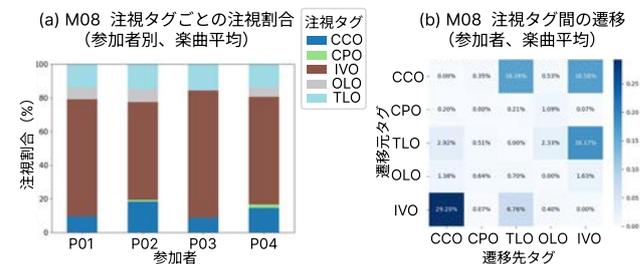


図 8 tool-based 型のモードの注視割合と遷移特徴 (IVO を中心とした注視行動)

さらに, 長時間表示型の *display duration* プリミティブと *finger-based gestures* 型の *user interaction* プリミティブも併せもつモードのユーザ観察データを分析すると, 2 種のプリミティブの注視行動に関する誘発性が異なる

ことが示唆された。前節で示したように、長時間表示型 *display duration* プリミティブでは、視線が散在しやすくなる傾向 (TLO や OLO への注視が増える) があったが、finger-based gestures 型の *user interaction* プリミティブを併せもつモードではこの傾向が見られず、視線は歌詞そのもの (CCO や CPO) に集中していた。たとえば、M23 や M24 では、注視遷移傾向は短時間表示の M01 とほぼ同様であった (図 7)。このことから、長時間表示型の *display duration* と finger-based gestures 型の *user interaction* プリミティブを併せもつモードでは、finger-based gestures 型の効果が *display duration* の効果を抑制するほど強く作用し、他の要素よりも優先的にユーザの注視行動を制御することが示唆された。

効果の抑制と制御という観点から、上述の議論をもとに各プリミティブの影響の大きさを整理すると、概ね以下の順に効果の強弱の関係があると示唆された。

- (1) 文字の大きさ (*scaling character*)
- (2) ユーザインタラクション (*user interaction: finger-based gestures*)
- (3) 文字の空間配置 (*display positioning*)
- (4) 文字の表示継続時間 (*display duration*)

5. 考察: インタラクション体験の構造

各注視タグの注視割合を通して見えてきたのは、インタラクションプリミティブがそれぞれ異なるレベルでユーザ体験を形成し、さらに複数のプリミティブが相互に作用しているという構造である。

scaling character は、体験の視覚的没入と身体的探索の度合いを決定する要素であると示唆された。*user interaction* が finger-based gestures であると、体験の能動性を高める身体的プリミティブとして機能し、*display duration* の効果を上書きすることが示唆された。*display positioning* は、体験の統制度と個人差を規定し、relative で統一的、absolute で探索的体験を誘発することが示唆された。*display duration* は、体験の時間的テンポをデザインするプリミティブである。表示時間が短いときには発声中の歌詞を追う集中的な体験が、長いときには歌唱の流れを離れて歌詞を見返す探索的な体験が生じることが示唆された。

これらのプリミティブの影響は独立ではなく、階層的・加算的に作用している。

6. おわりに

本稿では、MR 環境における歌詞体験を構成する 10 種のインタラクションプリミティブを提案し、うち 4 種のプリミティブについて、ユーザの注視行動に与える影響を既存のユーザ体験データの定量分析により明らかにした。

今後は、今回の分析では確認できなかった残り 6 種のインタラクションプリミティブの体験への効果と影響力につ

いて調査すると共に、MR 環境における、音楽と同期した文字表現を擁するインタラクション体験のデザインについての全体構造を明らかにすることを目指す。

参考文献

- [1] Kato, J. and Goto, M.: Lyric App Framework: A Web-Based Framework for Developing Interactive Lyric-Driven Musical Applications, *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, New York, NY, USA, Association for Computing Machinery, (online), DOI: 10.1145/3544548.3580931 (2023).
- [2] Ford, S., Forlizzi, J. and Ishizaki, S.: Kinetic typography: issues in time-based presentation of text, *CHI '97 extended abstracts on Human factors in computing systems looking to the future - CHI '97*, Atlanta, Georgia, ACM Press, p. 269 (online), DOI: 10.1145/1120212.1120387 (1997).
- [3] Kato, J., Nakano, T. and Goto, M.: TextAlive: Integrated Design Environment for Kinetic Typography, *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, New York, NY, USA, ACM, pp. 3403–3412 (online), DOI: 10.1145/2702123.2702140 (2015).
- [4] 横濱拓樹, 柏木敏郎, 山本恭裕, 中小路久美代: MR を用いた読書環境における文字の表現方式と詩の印象の違いに関する考察, *Human Interface Symposium 2022*, pp. 75–80 (2022).
- [5] Miyagawa, S., Koyama, Y., Kato, J., Goto, M. and Morishima, S.: Placing Music in Space: A Study on Music Appreciation with Spatial Mapping, *Proceedings of the 2018 ACM Conference Companion Publication on Designing Interactive Systems*, Hong Kong China, ACM, pp. 39–43 (online), DOI: 10.1145/3197391.3205409 (2018).
- [6] Jeon, J. and Lee, C. H.: SoundMist: Novel Interface for Spatial Auditory Experience, *Adjunct Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, San Francisco CA USA, ACM, pp. 1–3 (online), DOI: 10.1145/3586182.3616622 (2023).
- [7] Jeon, J. and Lee, C. H.: Sprayable Sound: Exploring the Experiential and Design Potential of Physically Spraying Sound Interaction, *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, Yokohama Japan, ACM, pp. 1–17 (online), DOI: 10.1145/3706598.3713786 (2025).
- [8] 小島颯英, 中小路久美代, 加藤 淳: MR 環境 LyricBathe におけるインタラクティブな歌詞表現のデザインスペースの創出, 信学技報 HCS2024-79, HIP2024-80 (2025-01), 茨木, 大阪, 日本, 電子情報通信学会, pp. 78–83 (2025).
- [9] 小島颯英, 中小路久美代, 加藤 淳, 後藤真孝: 歌詞没入型身体的音楽体験を可能とする光学透過式 MR 環境 LyricBathe, エンタテインメントコンピューティングシンポジウム 2025 論文集, Vol. 2025, 情報処理学会, pp. 28–37 (2025).
- [10] Yamamoto, Y., Nakakoji, K. and Takashima, A.: The Landscape of Time-Based Visual Presentation Primitives for Richer Video Experience, *Human-Computer Interaction - INTERACT 2005* (Costabile, M. F. and Paternò, F., eds.), Berlin, Heidelberg, Springer Berlin Heidelberg, pp. 795–808 (2005).
- [11] Suno, Inc.: Suno. <https://suno.com/home>, [2025-06-12].