

# League of Legends における実況解説音声の解析を通じたイベント発生予測

日高萌絵（指導教員：土田修平）

## 1 はじめに

近年、eスポーツ市場は目覚ましい成長を遂げており、日本国内においてもその注目度は高まっている。動画視聴や試合観戦の経験者は増加の一途を辿っており<sup>\*1</sup>、特に Multiplayer Online Battle Arena (MOBA) のようなチーム対戦型ゲームは、その戦略性と競技性の高さから、eスポーツシーンにおいて大きな人気を集めている。代表的なタイトルである League of Legends (以下 LoL) は、世界中で数百万人のプレイヤーと観戦者を擁し大規模な国際大会が開催されるなど、eスポーツの隆盛を牽引する存在となっている。<sup>\*2</sup>

しかしながら、eスポーツ観戦、特に LoL のような複雑なゲームにおいては初心者が直面する課題がある。画面上にはアイテムやキャラクターのステータスなど多くの情報が同時に表示され、重要な情報を瞬時に判断することが難しい上に、試合展開の理解には専門用語や戦略概念の前提知識が必要である。eスポーツに関する興味関心においては日常生活でのゲームプレイや実況動画の視聴経験の有無が関係しており<sup>\*3</sup>、初心者でも気軽に観戦を楽しめる環境の構築が急務となっていると考える。

これらの課題に対して、試合の状況や興奮を効果的に伝える重要な要素である実況解説音声に焦点を当てた。実況解説は専門的な情報を分かりやすく伝え、視聴者の感情的共感を促進することで、観戦体験をより豊かなものにする役割を果たしていると考えられる。

そこで本研究では、LoL の公式試合における日本語実況解説音声から自然言語処理技術を用いて試合状況に関する情報を抽出することを試みる。これにより、試合の重要な出来事（キル、タワー破壊、オブジェクト獲得など）を自動的に検出し、初心者向けに分かりやすく提示するシステムの構築を目指す。

## 2 関連研究

野球のような一般的なスポーツにおいては実況の音声情報の分析に関する研究の先例が挙げられる。佐藤ら [1] は映像がなくてもスポーツを楽しむ人に向けて映像から映像中の出来事や文字情報から既存の実況を補完する解説情報生成システムの構築に取り組んでいる。大久保ら [3] はスポーツ映像における行動認識において従来の画像特徴量のみではなく実況音声テキスト特徴量を加えた手法を作成し、実況音声テキスト特徴量の有効性と汎用性を明らかにしている。また、森本ら [2] は、LoL 試合の統計データの分析から対戦成績向上のためのパフォーマンス指標を明らかにしている。平ら [4] は、LoL のゲーム中の出来事から各プレイヤーの活躍度を算出しそれらを画面上に可視化するインタフェースに取り組んでいる。本研究では e スポーツの実況音声テキストから対戦成績に関わるパフォーマンスやイベントを抽出できるかについて調査を行った。

\*1 <https://jesu.or.jp/contents/news/news-231225/8>

\*2 <https://fistbump-news.jp/article/2024/11/21/460.html>

\*3 <https://biz.loyalty.co.jp/report/043/>「ロイヤリティマーケティング」調べ

## 3 分析対象

本研究で扱う LoL とは 5 対 5 で戦うチームゲームであり、プレイヤーは「チャンピオン」と呼ばれるキャラクターを操作し、マップ上のエピックモンスターや相手チームのチャンピオンを倒すことで得た経験値やお金で自身のキャラクターを強化しながら、相手陣地内の「タワー」を壊し最終的に「ネクサス」と呼ばれる相手チーム本拠地のオブジェクトを壊すことでそのプレイヤーが所属するチームが勝利となる。今回、公式試合のひとつである 2024 年の World Championship の日本語配信の実況解説の音声情報を分析した。決勝の試合を日本語の公式配信<sup>\*4</sup>から動画を取得し、また統計サイトである OP.GG Esports<sup>\*5</sup>と Games Of Legends<sup>\*6</sup>から試合の詳細なデータを取得した。

## 4 データ分析の手法とその結果について

### 4.1 試合中の実況音声データに基づく品詞出現頻度と頻出単語の差分析

はじめにゲーム内の重要なイベント発生時における実況解説の言語的特徴を分析することで、イベントの種類や試合の状況に応じた解説のスタイルや頻出する品詞の傾向を明らかにすることを目指した。

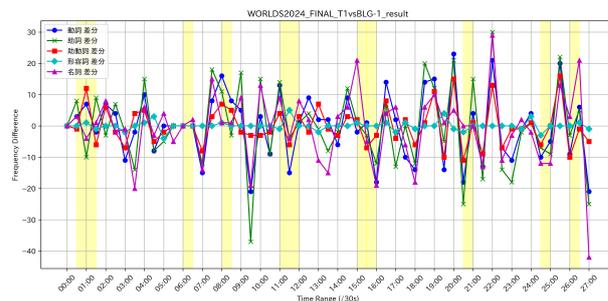


図 1 1 ゲーム中の各 30 秒単位の時間ブロックにおける品詞の出現数の差分

取得した動画から OpenAI Whisper を用いて動画の音声文字起こししたものに形態素解析を行い、動詞・助動詞・助数詞・形容詞の 4 つの品詞の出現回数の差分を時間軸に沿ってプロットしたものが図 3 である。黄色で着色されている時間帯は、ゲームの主な転換点であるファーストキル・ファーストタワー・ヴォイドグラフ/リフトヘラルド/ドラゴン/バロンナッシュアの 4 つの中立モンスターの獲得、の 6 種類のイベントが発生した時間帯である。動詞・名詞・助詞は特にイベント発生の時間帯に出現頻度が増えており、それ以外の時間帯では減っていることが読み取れる。イベントが発生している時間は必然と実行解説の発言量も多くなるためだと考えられる。次に、頻出単語上位 5 つを全ての品詞・名詞のみ・動詞のみを対象にして行った。

\*4 <https://www.twitch.tv/videos/2291754243>

\*5 <https://esports.op.gg>

\*6 <https://gol.gg/esports/home/>

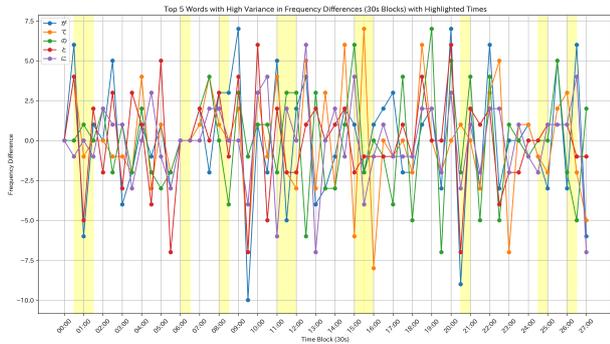


図2 各30秒単位の時間ブロックにおける全品詞の出現頻度上位5つの差分

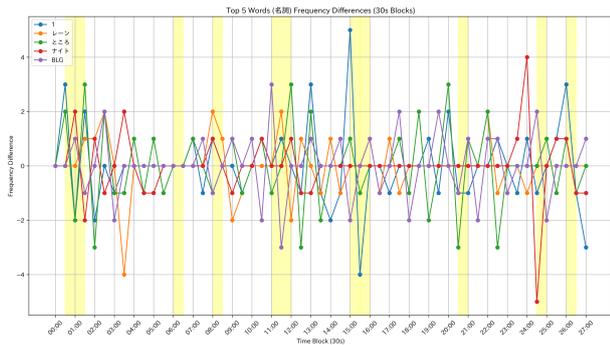


図3 各30秒単位の時間ブロックにおける名詞の出現頻度上位5つの差分

全体では助詞の出現頻度が高く、名詞に関しては特にイベント発生時とそれ以外での差分がわかりやすく反映されていることがわかる。

#### 4.2 機械学習を用いたイベント発生予測モデルの作成

次に、イベント発生時において単語の出現頻度に変化することから、頻出頻度からのイベント検知に取り組んだ。3.1で使用したイベントのみだとイベント発生時のラベル数が少ないため、過学習を避けるために各チームが相手チームのチャンピオンを倒した時間もイベントに含めた。これらのイベント発生時の時間帯を event 値=1、それ以外を0とし、15秒単位での各品詞の出現数の特徴量とし標準化した上でSVMを用いて分類モデルの構築と評価を行った結果が表1である。

表1 1試合における分類モデルの評価結果

評価結果	適合率	再現率	f1-score (調和平均)	サンプル数
クラス0	0.75	0.64	0.69	14
クラス1	0.44	0.57	0.50	7
正解率			0.62	21
マクロ平均	0.60	0.61	0.60	21
重み付き平均	0.65	0.62	0.63	21

クラス1すなわちイベント発生時の検知の成功割合が0.5を切っていて精度が低くなっている。表1の段階では訓練データのラベルのクラス分布がクラス0=53、クラス1=30となっており不均衡さが目立つため、次に訓練データにSMOTEを適用しオーバーサンプリングを行った。訓練データのラベルのクラス分布がク

ラス0 = クラス1 = 53の状態再度分類モデルの評価を行ったがSMOTE適用前との有意な差は見られなかった。

また、5秒単位での各品詞の出現数を特徴量としたところ精度が大幅に下がったため、5秒単位での出現数とその変化量の和を特徴量にし、同様に再度分類モデルの評価を行ったがこちらも表1より低い精度となった。この時点での特徴量とラベルにおいて5分割の交差検証を行った結果が表2である。

表2 5分割交差検証の結果

	正解率	f1-score (調和平均)
スコア	[0.65 0.48 0.61 0.57 0.53]	[0.49 0.42 0.46 0.30 0.44]
平均	0.57	0.42
標準偏差	0.059	0.064

適合率、F1スコアともに標準偏差の数値がやや大きいことから分割箇所によってばらつきがあること、また4つ目のテストではどちらのスコアもパフォーマンスが落ちている事がわかった。したがって1試合のなかで訓練データとテストデータを取る場合、その分け方の詳しい検証も今後必要になると考えられる。

### 5 今後の展望

本研究ではLoLの公式試合の実況解説の音声情報からの試合状況の推察として、形態素解析による品詞ごとの出現頻度とイベント発生との関係の分析とそこから予測モデルの構築に取り組んだ。イベント発生時に品詞の出現頻度が増加する傾向が強いことはわかったが、予測モデルの構築に関してはクラス1の適合率を50%以上にすることができなかった。

今回のSVMを用いた予測モデルの構築では1ゲームの中で訓練データとテストデータを分割し、また特徴量は時間ごとの各品詞出現数で固定した。今後の課題として、訓練データとテストデータに違うゲームを割り当てることや、特徴量を品詞の組み合わせにするなど変更させることが挙げられる。

### 参考文献

- [1] 清山 信正 今井 篤 山田 一郎 佐藤 庄衛, 熊野 正. スポーツ中継のリアルタイムデータからの解説音声自動生成実験. **研究報告音声言語情報処理 (SLP)**, 2016-SLP-113(6):1-6, 2016.
- [2] 足立 潤治 森本 秀斗, 河瀬 彰宏. Lolにおけるチームの成績向上に関連するパフォーマンス指標の分析. **第86回全国大会講演論文集**, 2024(1):823-824, 2024.
- [3] 西山 裕之 大久保 深, 秦野 亮. スポーツ放送映像におけるマルチモーダル行動認識-画像特徴量と実況音声テキスト特徴量の統合-. **第86回全国大会講演論文集**, 2024(1):751-752, 2024.
- [4] 梶並 知記 平 優毅. League of legendsにおけるキャラクターの活躍度の可視化システム. **研究報告グループウェアとネットワークサービス (GN)**, 2023-GN-118(29):1-8, 2023.