

音声コミュニケーションにおける叫び声*

☆菊地佑樹, 森大毅 (宇都宮大)

1 はじめに

興奮, 驚きなどが叫び声として溢れ出ることがある。これは, 危機的状況に置かれた時に発する悲鳴だけではない。例えば, 気が許せる相手同士でスポーツ観戦をしている時に思わず叫ぶことがある。

叫び声に関する過去の研究の多くは防犯を目的としたものである [1][2]。一方, 緊急時でない日常のコミュニケーション場面で使用される叫び声は, 笑い声と同様, 音声コミュニケーションを豊かにするノンバーバル行動であるが, 研究事例はほとんどない。

本研究は, 音声コミュニケーションにおける叫び声の音響的特性, 及びそれが社会的相互行為の中で果たす役割を明らかにすることを目的としている。

2 コミュニケーションにおける叫び声

2.1 叫び声の定義

感情表出系感動詞は話者の認知的状態の変化を伝えるシグナルである [3]。感情表出系感動詞の中には, 疑問を感じたときに発する「え」や何かに気づいたときに発する「あ」など, 書き起こしが容易で単語として様式化されたものもある。しかし, 予想外の突然のイベントに対して表出される声の中には, 韻律あるいは声質の点で特異で, 「あ」に近い音ではあってももはや言語音とは言い難いものもある。本研究では, 感情表出系感動詞の中でも様式化の程度が低く, 話者の制御下にある度合いが低い (思わず発せられた) ものを叫び声としている。

2.2 叫び声の同定

対話音声コーパスに対して叫び声の同定を行った。対象としたコーパスは四コマ漫画並べ替え課題を解く様子を収録した宇都宮大学パラ言語情報研究向け音声対話データベース (UADB)[4], オンラインゲーム中の音声チャットを収録した感情評定値付きオンラインゲームチャットコーパス (OGVC)[5], 自由会話を収録した千葉大学三人会話コーパス (Chiba3Party)[6] である。叫び声の同定は第一著者が行った。Table 1 に対象とした発話数と叫び声の数を示す。UADB, OGVC, Chiba3Party における叫び声は多くないことが分かる。これらの中では OGVC における叫び声の頻度が高かった。これは, OGVC がゲーム中の音声チャットを収録しており, 他の 2 つのコーパスに比べて話者

Table 1 既存コーパスに含まれる叫び声

	UADB	OGVC	Chiba3Party
発話数	4840	8662	7599
叫び声数	4	53	4
叫び声頻度 (個/分)	0.0333	0.0803	0.0111

が予期しないような事象が頻繁に発生したためと考えられる。

3 叫び声研究のためのコーパス構築

既存のコーパスには多くの叫び声が含まれていないことが分かった。そこで会話中に自然に表出される叫び声をより多く得るために, アクション性の高いゲームをプレイ中の二国会話音声の収録を新たに行った。使用したゲームタイトルはアクション性が高いジャンルである FPS (First Person Shooter) のゲームである「OverWatch」と, 対戦型サッカーゲームである「FIFA16」である。ヘッドセットマイクによる音声のほか, 対話参加者の上半身動画像及びゲーム画面を同時に収録した。このコーパスを以後アクションゲーム音声コミュニケーションコーパス (AGSC) と呼ぶ。

Fig. 1 に収録を行った環境を示す。実験参加者は親近性の高い同性の 2 人組の大学生 8 組 (合計 16 名, 内男性 4 名, 女性 12 名) である。音声収録は宇都宮大学学内の対話収録用防音室で行った。

収録したデータのうち, 話者 2 名ずつ計 4 名 (男性 2 名, 女性 2 名) の各 10 分, 計 40 分 503 発話のデータに対し, 2 名の作業員による叫び声アノテーションを実施し, 一致度を調べた。アノテーションは著者 2 名で行った。各発話に対し, 叫び声の存在有無のアノテーションの一致度を Cohen の κ 係数により求めたところ, $\kappa = 0.663$ となった。

503 発話中, 第一著者により同定された叫び声は 45 個であった。叫び声頻度は 1.13 個/分であり, これは Fig. 1 に示した OGVC の叫び声頻度の約 14 倍である。このことからアクションゲームを利用した会話の収録により, 叫び声を効率的に収集することができたといえる。

*Shout in speech communication. by KIKUCHI, Yuki and MORI, Hiroki (Utsunomiya University)

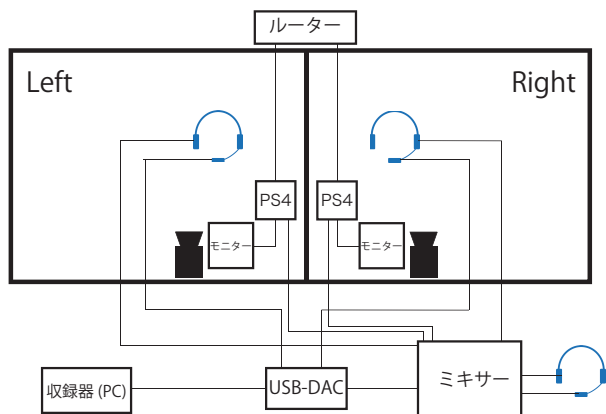


Fig. 1 収録環境

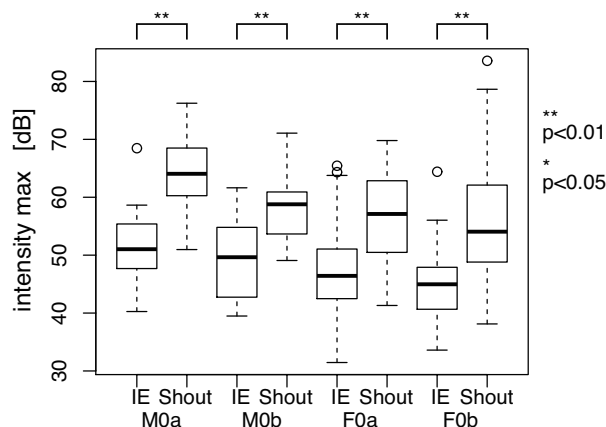


Fig. 3 強度最大値

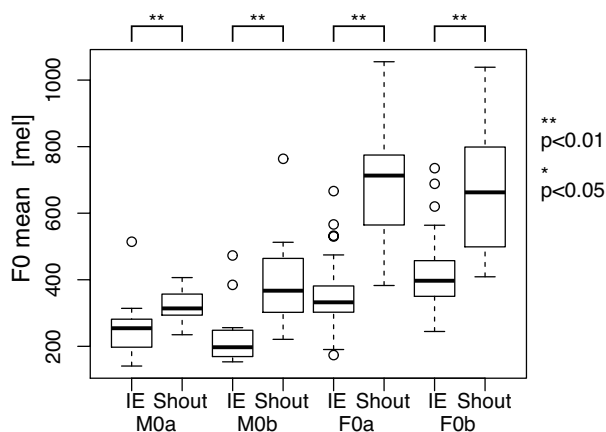


Fig. 2 F0 平均値

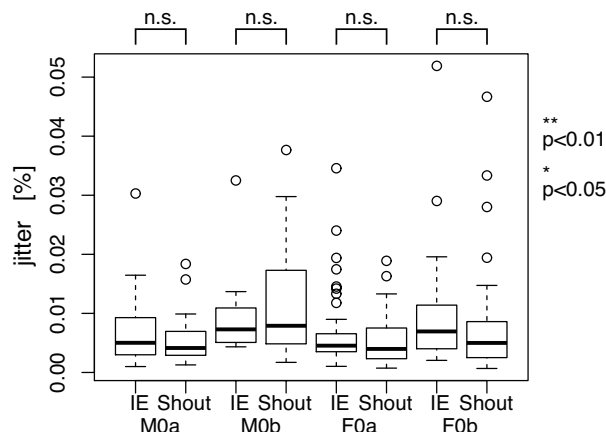


Fig. 4 ジッター

4 音響分析

叫び声の音響的特徴を調べた。対象とした発話は話者 M0a, M0b, F0a, F0b (男性 2 名, 女性 2 名) の感情表出系感動詞 192 個と叫び声 192 個の計 384 個である。F0 平均値, 強度最大値, 及びジッターの分析結果を Fig. 2, 3, 4 に示す。これらの図より, 叫び声は感情表出系感動詞より F0 平均値と強度最大値が大きいことが分かる。感情表出系感動詞と叫び声の F0 平均値の中央値の差が男性では約 150 mel, 女性では 200 mel 以上あり, 強度最大値の差に関しても男性女性ともに 10 dB ほどあることが読み取れる。一方, ジッターには感情表出系感動詞と叫び声の間に, 有意差は認められなかった。

5 おわりに

コミュニケーションにおける叫び声の特性を明らかにするため, アクションゲームに基づくコーパスを構築した。また, 音響的特徴の初歩的な分析を行った。その結果, 叫び声は発声の際に通常の発声より高く大きな声で表出されていることが分かった。

今後は, AGSC のアノテーション作業を進め叫び声と感情表出系感動詞の他の音響的特徴量分析を行う。また, 叫び声の音響モデルを作成し叫び声の合成を行っていきたい。

謝辞 本研究は JSPS 科研費 26280100 の助成を受けている。

参考文献

- [1] 柿野他, 音講論 (秋), 1-P-32d, 2013.
- [2] 近藤他, 情報処理学会研究報告, Vol. SLP-113, No. 2, 2016.
- [3] 森, 音講論 (春), 2-Q-43, 2015.
- [4] Mori et al., Speech Communication, Vol. 53, pp. 36–50, 2011.
- [5] 有本他, 音講論 (秋), 1-P-46a, 2013.
- [6] Den and Enomoto, in *Conversational informatics: An engineering approach*, 307–330, 2007.