

音声からの知覚による感情カテゴリと感情次元の関係*

○有本泰子 (芝浦工大), 森大毅 (宇都宮大)

1 はじめに

大規模音声コーパスを必要とする音声認識や感情認識などの人間と機械との対話の円滑化を念頭にした研究では、複数のコーパスを併用することが求められる。しかし、コーパスごとに独自基準で感情ラベリングを行っているため、感情ラベルをコーパス間で等価とみなすことができず、現状では複数コーパスの併用は不可能である。複数コーパスを用いて研究を進めるためには、音声コーパス間で共通化した感情ラベルが必要であり、音声に対する感情ラベルの標準化が求められる。本研究では、音声に自発的に表出する感情を対象に感情ラベルの標準化を目指し、まずは感情ラベルの多様性の問題に着目し、代表的な感情ラベルである感情次元と感情カテゴリの関係について検証を行った。本報告では、既存のふたつの感情音声コーパス(感情評定値付きオンラインゲーム音声チャットコーパス(OGVC) [1] および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(UUDB) [2])に付与されている異なる基準のラベルを相互に付与し、感情次元により構成された感情空間に感情カテゴリを配置することで、異なる感情ラベル間の関係について検証する。

2 クロスコーパス感情ラベリング

2.1 音声資料

本研究では、異なる基準の感情ラベルが付与されているOGVC[1]とUUDB[2]を用いる。

OGVCには、Massively Multi-player Online Role-Playing Game (MMORPG)をプレイ中に音声チャットを用いてプレイヤー間で対話を行わせた自発対話音声が取録されている。音声提供者はオンラインゲームの経験のある13名の日本人(女性4名, 男性9名)であり、9114発話が取録されている。このうち、音声の振幅レベルが小さい、または、笑いや息や咳を示す転記用タグが付与されている発話を除いた6578発話を本研究に用いた。付与されている感情ラベルはPlutchikの基本8感情(喜び(JOY), 受容(ACC), 恐れ(FEA), 驚き(SUR), 悲しみ(SAD), 嫌悪(DIS), 怒り(ANG), 期待(ANT)) [3]に加えて感情が表出していない平静(NEU)と、これらのどの感情にも分類できないその他(OTH)の10種類である。

UUDB[2]は「4コマまんが並べ換え課題」を取録したコーパスである。音声提供者は日本人大学生7ペア

Table 1 Result of emotional category labeling.

Emotion	OGVC	UUDB	Total
JOY	438	259	697
ACC	623	1030	1653
FEA	282	94	376
SUR	313	120	433
SAD	488	331	819
DIS	970	406	1376
ANG	128	39	167
ANT	186	59	245
NEU	18	13	31
Total	3446	2351	5797

(女性12名, 男性2名)であり、4840発話が取録されている。付与されている感情ラベルはRusselの円環モデル[4]を含む6つの感情次元(快-不快(pleasant-unpleasant), 覚醒-睡眠(aroused-sleepy), 支配-服従(dominant-submissive), 信頼-不信(credible-doubtful), 関心-無関心(interested-indifferent), 肯定的-否定的(positive-negative))である。各感情次元を1から7まで(4が中立)の7段階で評価されている。

2.2 感情ラベリング

本研究で用いた感情ラベルはOGVCで採用している感情カテゴリラベル(5段階(1-5)の強度を含む)とUUDBで採用している7段階(1-7)の感情次元ラベルの二つである。この二種類の感情ラベルをふたつのコーパスに相互に付与した。安定した感情ラベルを得るため、まずは被験者スクリーニングのための聴取実験を行った。実験参加者はコーパスの話者と同年代の10名(男性4名, 女性6名, 平均年齢: 21.8歳 ($SD = 0.84$))である。評価対象音声は108発話(OGVC: 9感情×3段階(弱・中・強)×各2発話=54発話, UUDB: 3段階(弱・中・強)×各18発話=54発話)である。一人の被験者に対し、発話の提示順序を変えて8回試行を行った。

被験者スクリーニングにより、3名(男性1名, 女性2名)のラベラーを選定し、本実験を行った。ラベラーの平均年齢は22歳 ($SD = 0.82$)である。評価対象の音声は11,418発話(OGVC:6578発話, UUDB:4840発話)で、聴取回数は1回とした。付与したラベルは被験者スクリーニングと同じである。

感情カテゴリラベリングの結果、3名のラベラーのうち2名が一致した感情をその発話の感情カテゴリ

*Relationship between emotional category and emotional dimension perceived from speech. by ARIMOTO, Yoshiko (Shibaura Institute of Technology), MORI, Hiroki (Utsunomiya university))

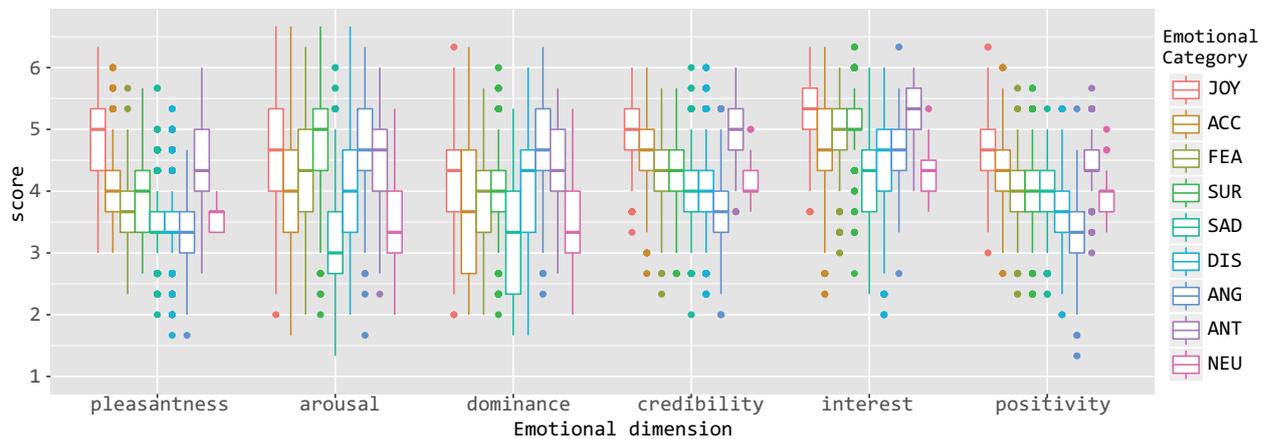


Fig. 1 Distribution of each emotional category on emotional dimensions.

ラベルとみなす。各感情カテゴリの発話数を Table 1 に示す。

3 感情次元における感情カテゴリ

Fig. 1 に感情カテゴリごとに感情次元ラベルの分布を示す。感情次元ごとに感情カテゴリを要因とした Kruskal-Wallis 検定を行った。その結果、いずれの感情次元においても有意差を示した ($p < 0.001$)。感情次元内で感情カテゴリ間に有意差を示す組合せを調べるため、Bonferroni 法による多重比較を行った。

pleasantness 次元では、JOY の発話が最も値の高いところ (快状態) に分布しており、ANG の発話が最も低いところ (不快状態) で分布している。他の 8 感情と有意な差を示した ($p < 0.05$) のは JOY・ACC・ANG・ANT であった。Fig. 1 から、JOY > ANT > ACC > ANG の順に pleasantness 次元の値が高いことが分かる。arousal 次元では、多くの感情カテゴリが広範囲に渡って分布している。その中で SAD が最も低いところ (睡眠状態) で分布している。SAD は NEU 以外の 7 感情と有意な差を示し ($p < 0.05$)、arousal 次元において SAD のみが他の感情と異なることが分かる。dominance 次元では、ACC と SAD が NEU 以外の 7 感情と有意差を示した ($p < 0.05$)。また、ANG と ANT は相互に有意差を示さなかったが、他の 7 感情と有意差を示した ($p < 0.05$)。Fig. 1 から、ANG = ANT > ACC > SAD の順に dominance 次元の値が高いことが分かる。credibility 次元では、ACC と ANG が他の 8 感情と有意な差を示した ($p < 0.05$)。さらに、ANT と JOY は相互に有意差を示さなかったが、他の 7 感情と有意差を示した ($p < 0.05$)。また、Fig. 1 から、JOY = ANT > ACC > FEA = SUR > SAD = DIS > ANG の順に credibility 次元の値が高いことが分かる。interest 次元では、JOY と ANT は相互に有意差を示さなかったが、他の 7 感情と有意差を示した ($p < 0.05$)。また、SAD と DIS は NEU 以外の 7 感情と有意差を示した ($p < 0.05$)。また、

Fig. 1 から、JOY = ANT > SAD = DIS の順に interest 次元の値が高い。positivity 次元では、JOY と ANG が他の 8 感情と有意な差を示した ($p < 0.05$)。さらに、SUR と DIS は NEU 以外の 7 感情と有意差を示した ($p < 0.05$)。さらに、ACC と ANT 相互に有意差を示さなかったが、他の 7 感情と有意差を示した ($p < 0.05$)。Fig. 1 から、JOY > ACC = ANT > SUR = DIS > ANG の順に positivity 次元の値が高くなっている。

4 感情空間上における感情カテゴリ

Fig. 2 および Fig. 3 に、pleasantness, arousal 次元を用いた 2 次元の感情空間上に発話の分布をマッピングした結果を示す。Fig. 2 は pleasantness-arousal 空間、Fig. 3 は pleasantness-dominance 空間である。図中のマーカーの大きさはその位置に分布する発話数を示す。オレンジ色のマーカーは OGVC の発話、緑色のマーカーは UUDB の発話である。

Fig. 2 は Russell が提案した円環モデル [4] と同じ感情空間である。Russell が提案した感情空間内に配置される感情カテゴリと比較すると、JOY は高覚醒度で快状態であるため、Russell の提案と一致する。さらに、ANG は高覚醒度で不快状態、SAD は睡眠状態で不快状態、SUR が高覚醒度であることも、Russell の提案と矛盾しない。これにより、本研究の結果はほぼ妥当であると考えられる。NEU については pleasantness 次元では中立 (4) 付近に位置しているのに対し、arousal 次元では広い範囲に分布している。さらに、dominance 次元の NEU を確認しても、広範囲に分布している。この結果は、感情を含まない平静状態の発話というものが pleasantness の状態が中立であることであって、他の感情次元は中立である必要はないことを示唆している。

また、Fig. 2 および Fig. 3 より、異なるコーパスが同じ感情カテゴリの発話を含んでいたとしても、感情空間上ではその分布が異なることが示された。Fig. 2

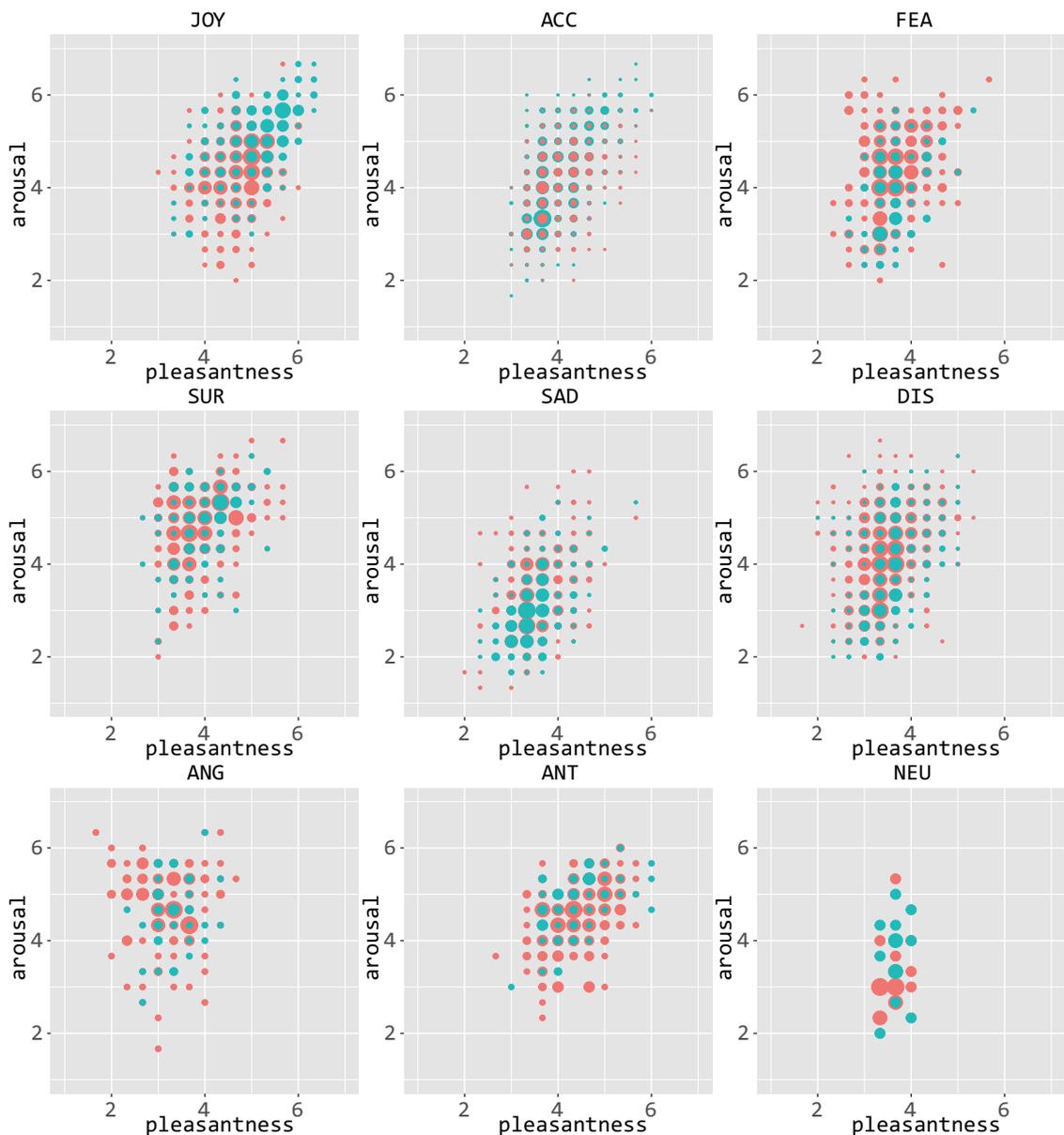


Fig. 2 Mapping each emotional category into pleasantness–arousal dimension. Orange markers represent the speech samples in OGVC and green markers represent the speech samples in UUDB.

のJOYの図では、UUDBの発話の方がOGVCよりも右上に多くの発話が分布しており、UUDBのJOYの発話が快状態と高覚醒度をともに強く知覚させることが分かる。同様に、ANGの図では、OGVCの発話の方がUUDBよりも左上に多く分布しており、OGVCのANGの発話が不快状態と高覚醒度をともに強く知覚させることが分かる。Fig. 3のACCおよびSADの図では、UUDBの発話の方がOGVCよりも下に多く分布しており、UUDBのACCの発話が服従状態を強く知覚させることが分かる。さらにDISにおいては、UUDBの発話が2か所に分かれて分布していることが分かる。ひとつはOGVCと同様に中央付近に分布し（やや支配状態）、もうひとつは下に分布している（服従状態）。同一のコーパス内で異なる種類のDISの発話が存在することが示される。

5 おわりに

音声に自発的に表出する感情を対象に感情ラベルの標準化を目指し、まずは感情ラベルの多様性の問題に着目し、代表的な感情ラベルである感情次元と感情カテゴリの関係について検証を行った。既存のふたつの感情音声コーパスに付与されている異なる基準のラベルを相互に付与し、感情次元により構成された感情空間に感情カテゴリを配置することで、異なる感情ラベル間の関係について検証を行った。その結果、感情空間上に配置された感情カテゴリは先行研究と概ね同じ結果となったものの、平静の発話に関しては次元の両極に偏らなかったのはpleasantnessのみであり、平静の発話が快–不快次元以外ではかならずしも中立と知覚されないことが示唆された。さらに、

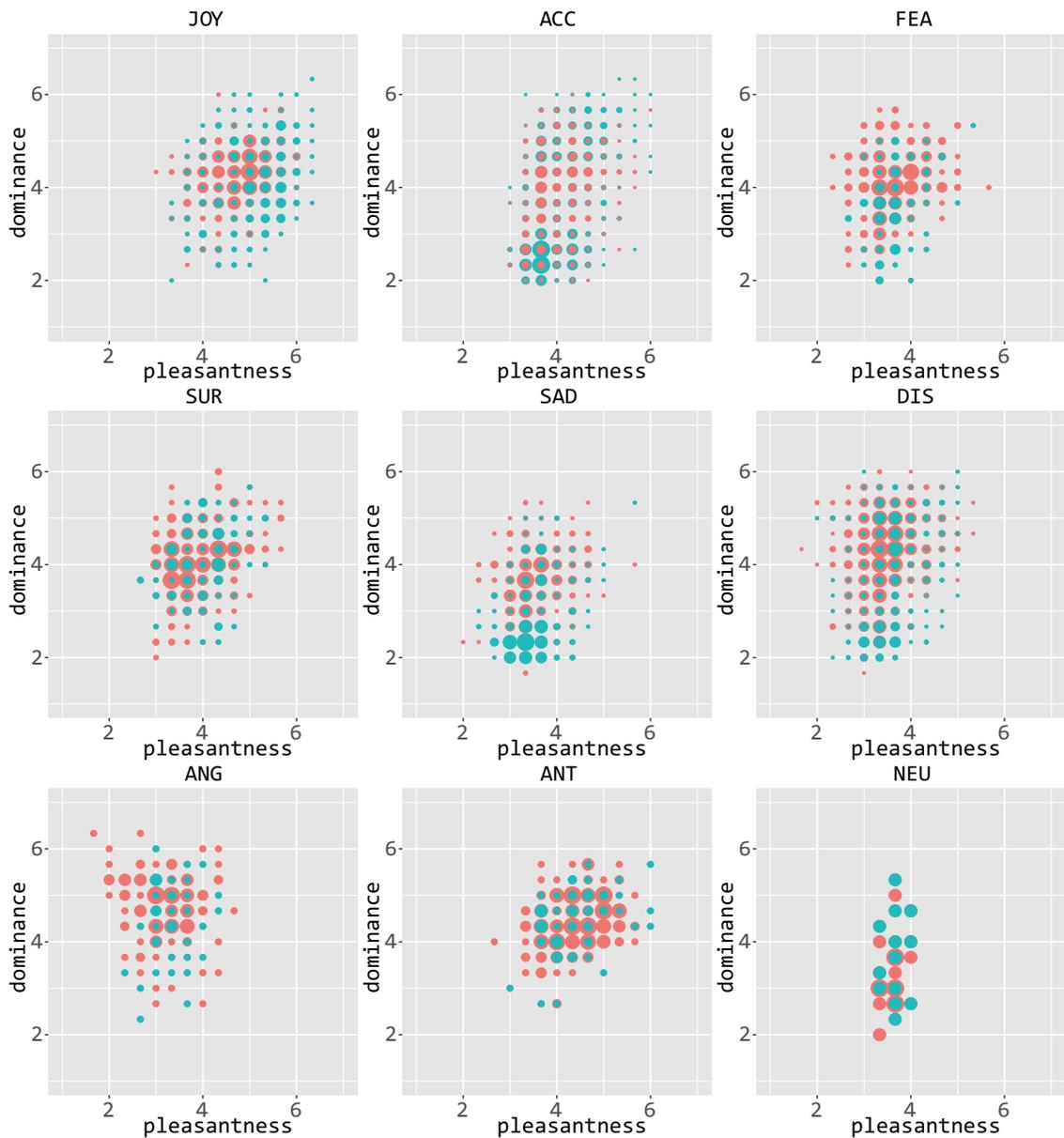


Fig. 3 Mapping each emotional category into pleasantness–dominance dimension. Orange markers represent the speech samples in OGVC and green markers represent the speech samples in UUDB.

感情空間上に配置された発話をコーパスごとに分類すると、同じ感情カテゴリを知覚させる発話であっても、コーパスごとに感情空間上の異なる位置に配置されることが示された。コーパスの対話ドメインによって感情の表現に差が生じている可能性がある。

今後は、発話の音響的特徴と感情次元・感情カテゴリの関係について精査し、音声によって知覚される妥当な感情次元について検証を重ねていく予定である。

謝辞 本研究は財団法人立石科学技術振興財団研究助成 (A) および JSPS 科研費 (26280100) の助成を受けた。

参考文献

- [1] Arimoto *et al.*, *Acoust. Sci. Tech.*, 33 (6), 359–369, 2012.
- [2] Mori *et al.*, *Speech Commun.*, 53(1), pp. 36–50, 2011.
- [3] Plutchik, “Emotions: A Psychoevolutionary Synthesis.” Harper & Row, 1980.
- [4] Russell, “Emotion: Theory, Reresearch, and Experience,” 4. 83–111, Academic Press, 1989.