

パラ言語情報の抽象次元評価における 被験者間の一致度と許容度の関係*

◎相澤宏, 森大毅, 粕谷英樹 (宇都宮大・工)

1 はじめに

高度な音声認識・音声合成の実現には、発話者の意図・心的態度・感情などのパラ言語情報についての深い理解が重要であるという意識が高まってきている。高度な音声認識・音声合成を実現するためには音声の持つパラ言語情報を定量的に表す必要があるが、まだその方法は確立していない。

我々は音声を被験者に聴取させ、発話中のパラ言語情報をどのように評価するか研究してきた[1]。その結果、パラ言語情報の評価が被験者間で一致するものも多いが、一致しない発話もあることがわかった。このことは、人間の持つ多様性を考えれば自然な結果であるが、評価が一致しないことが持つ意味はわかっていない。しかし被験者間の評価が一致しなくとも、被験者が互いの評価に対して「そういう評価もありうる」と考えるならば、評価が一致しなければならぬとまでは言えない。そこで今回は新たにパラ言語情報の評価に対する許容度という概念を導入する。どのような評価値でも許容度が高ければ、その発話のパラ言語情報は様々な値を取りうることになる。今回は被験者間のパラ言語情報の一致度と許容度の関係を聴取実験を通じて検討する。

2 パラ言語情報の記述

本研究では感情心理学の分野で多く用いられている以下の6軸の評価項目を使用して、対話におけるパラ言語情報を記述する。6軸の評価項目は快-不快と覚醒-睡眠を基本的な心理状態の評価項目とし、その他を対人関係、対人評価を考慮した評価項目で構成してある。

- 快-不快 発話者の気持ちや気分の良し悪し
- 覚醒-睡眠 発話者の心理状態の活発さ
- 支配-服従 発話者が相手とのコミュニケーションをリードしているか

- 信頼-不信 発話者が相手のことをどの程度信じているか
- 関心-無関心 発話者が相手や相手の発話に対してどの程度関心や興味があるか
- 肯定的-否定的 発話者が相手の発話をどの程度肯定的に評価しているか

3 実験方法

実験で使用する発話はパラ言語情報を多く含む一連の発話である。対話者は親しく話せる女性2組計4名で発話数は215である。被験者は正常な聴力を持つ20歳代の男女5名であり、実験の協力に対して謝金が支払われている。発話の再生は被験者自身がPC上の音楽ソフトウェアを操作して行い、同一のヘッドホンを用いて聴取する。

まず被験者自身にパラ言語情報を評価させる。被験者は、全発話を順に再生し各発話に対し6軸の評価項目を7段階で評価した。

次に許容の程度を求める実験を行う。被験者は Fig. 1のように提示された評価値に対して

- 同意できる (その通り)
- 許容できる (まあ良いのではないか)
- 許容できない (それは違う)

のいずれかを選択する。被験者には「ある人はこの発話に対して以下のように○を付けました。あなたはこの評価をどう思いますか。一番近いものをチェックしてください。」と教

No.1
ある人はこの発話に対して以下のように○を付けました。
あなたはこの評価をどう思いますか。一番近いものをチェックしてください。

不快 1—②—3—4—5—6—7 快
 同意できる (その通り) 許容できる (まあ良いのではないか) 許容できない (それは違う)

睡眠 1—②—3—4—5—6—7 覚醒
 同意できる (その通り) 許容できる (まあ良いのではないか) 許容できない (それは違う)

Fig. 1 許容の程度評価シート

* Relation between the degree of agreement and acceptability of paralinguistic information evaluation based on abstract dimensions. By AIZAWA Hiroshi, MORI Hiroki, KASUYA Hideki (Utsunomiya University)

示した。実験は同一の発話に対して7回行い、その中で提示する評価値が各項目1~7まで1回ずつ含まれるようにした。すなわち実験を全て行えば1~7までのパラ言語情報の評価値に対する許容の程度がそれぞれ求まる。許容の程度を求める発話はこれまでの研究[1]における被験者間の分散が大きい発話・小さい発話・中程度の発話の合計44発話とする。許容の程度を判定するにあたっては発話を複数回聴取可能とした。

4 実験結果

各発話の各項目で許容度を求めた。許容度とは評価項目ごとに1~7の評価値で「同意できる」もしくは「許容できる」とされた割合と定義する。許容度の平均値をTable1に示す。もし1~7の評価値のうち1つだけ許容されていれば、許容度は $1/7 \approx 14\%$ となる。許容度の平均値をTable1に示す。Table1よりどの項目でも許容度が大きいことがわかる。Table1の数値はどの評価項目も1~7の値のうち5つ程度が許容されていることを示しており、許容される範囲が広い。パラ言語情報の記述は1つの値だけでなく範囲を持って表す方法も考えられる。

次に許容度と一致度との相関係数を求めた。一致度 a は各ラベラの評価値が全体と一致した割合であり、次式で定義する。

$$a = \frac{|\{i | i \in S, |x_i - \bar{x}| \leq 0.5\}|}{|\{i | i \in S\}|}$$

ただし、被験者集合 S は[1]と今回の5名の被験者から成る。結果をTable2に示す。Table2より許容度と一致度の相関が弱いことがわかる。Table1で許容度の平均値が高いことと合わせて考えると、被験者間の一致度が高い低いに関わらず許容されていることがわかる。

次に被験者ごとの許容度をTable3に示す。Table3では特に#4が許容する割合が多いなど、被験者によって許容する傾向に違いがあることがわかる。これより他の被験者より許容する傾向にある被験者はパラ言語情報の評価に自信がない可能性が考えられる。今後、許容度をパラ言語情報を評価する被験者のスクリーニング基準として利用できないか検討したい。

次に被験者本人の評価値と提示した評価値

の許容の程度との関係を調べた。それぞれ「同意できる」を2, 「許容できる」を1, 「許容できない」を0と数値化した値を許容値と呼ぶ。被験者本人の評価値と提示した評価値の絶対値を求め、許容値との相関を算出した。その結果全ての評価項目で負の相関があった。このことから被験者本人の評価値から離れた評価値ほど許容されないことが確認できた。

5 まとめ

パラ言語情報の評価に対する許容度という概念を導入し、提示された評価値に対する許容度が持つ性質を5名の被験者による実験を通して調べた。その結果、被験者間でパラ言語情報の評価が一致する・しないに関わらず許容度は高かった。今後は許容度が高くなる原因や、1~7の評価値に対する許容の分布をより詳細に検討したい。

謝辞

有益な議論をしていただいた宇都宮大学国際学部 中村真 助教授に感謝する。

参考文献

- [1] 森大毅他, “対話音声のパラ言語情報ラベリングの安定性,” 日本音響学会誌 (掲載予定) .

Table 1 許容度の平均値

	許容度の平均値
快-不快	0.65
覚醒-睡眠	0.73
支配-服従	0.71
信頼-不信	0.63
関心-無関心	0.66
肯定的-否定的	0.68

Table 2 許容度と一致度の相関係数

	相関係数
快-不快	0.03
覚醒-睡眠	-0.24
支配-服従	-0.22
信頼-不信	-0.34
関心-無関心	0.03
肯定的-否定的	-0.11

Table 3 被験者ごとの許容度

被験者ID	#1	#2	#3	#4	#5	全体
同意できる or 許容できる	72%	59%	62%	82%	61%	67%
許容できない	28%	41%	38%	18%	39%	33%