

## 自然な対話における発話間ポーズ長と 知覚されるパラ言語情報との関係\*

○森 大毅 (宇都宮大)

### 1 はじめに

人間同士の対話における発話時期の調整には、多くの要因が関与している。話者交替や相槌の生起位置については、発話の言語的および韻律的特徴との関係が示されている [1, 2, 3]。また、先行発話から当該発話までのポーズ長に関しては、対話者間での同調傾向が指摘されている [4]。

これらに加え、話者の感情状態や態度なども、発話間ポーズ長に影響を与える要因となっていることが予想される。例えば、あまり自信のない発話の前のポーズ長は長くなることが考えられる。また、強く同意する発話では負のポーズ長、すなわち先行発話にオーバーラップした発話ともなり得る。

本論文は、親しい友人同士の表情豊かな対話における発話間ポーズ長が話者の感情状態などから受ける影響を明らかにすることを目的とする。

### 2 分析方法

以下の分析における用語を定義しておく。

**スラッシュ単位** 話し言葉における「文」に相当する単位。

**TRP** 移行適格場所 [5]。話し手のターンが終了し得る点。中断以外のスラッシュ単位の終わりはほぼ TRP に等しい。

**正常話者交替** TRP において生じた話者交替。TRP は予測可能であるため、2 者の発話がオーバーラップする場合もある。

**相槌** 相手の発話を促す「うん」「はい」などの発話。スラッシュ単位を構成せず、相手のスラッシュ単位の途中にしか現れない。

**割り込み** TRP 以外における相手話者の非相槌発話。話者交替が生じる場合と生じない場合とがある。

発話間ポーズ長の分析は、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) [6] に含まれる大学生 7 ペア (女性 12 名, 男性 2 名) による対話を対象とする。UUDB では、400 ms

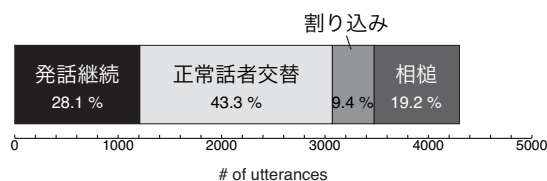


Fig. 1 Categorization of utterances.

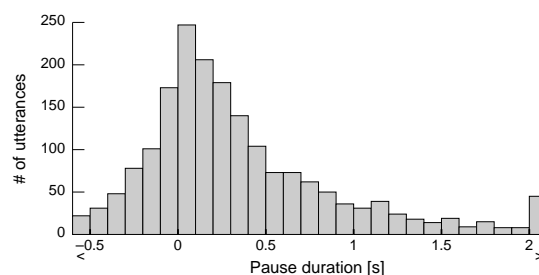


Fig. 2 Pause duration distribution.

以上の無音およびスラッシュ単位に挟まれた区間を発話単位とし、各発話から知覚される話者の感情状態が、「快-不快」「覚醒-睡眠」「支配-服従」「信頼-不信」「関心-無関心」「肯定的-否定的」の 6 抽象次元によりラベリングされている。ただし、感情状態のラベリングは各発話を独立に聴取して行われたため、ラベラは発話間ポーズを聴いていない。

さらに、著者 1 名による対話音声の聴取に基づき、各発話に対する先行発話の同定を行った。さらに、先行発話が相手話者の発話である場合、当該発話の開始点が先行発話の TRP であるか否かを、言語的および韻律的手がかりから総合的に判定した。この手づきにより、UUDB 中の各発話は Fig. 1 のように分類された。TRP において、先行話者の次発話開始と相手話者の発話開始が重複した場合は、発話継続と正常話者交替が同時に起こっているものとした。なお、先行発話が不明確な発話は分析対象から除いた。

今回は正常話者交替と分類された 1862 発話を対象とし、当該発話の開始時刻と先行発話の終了時刻との差を求めることで発話間ポーズ長を求めた。その度数分布図を Fig. 2 に示す。

\* Response latencies and perceived paralinguistic information of utterances in spontaneous dialogue.  
by MORI, Hiroki (Utsunomiya University)

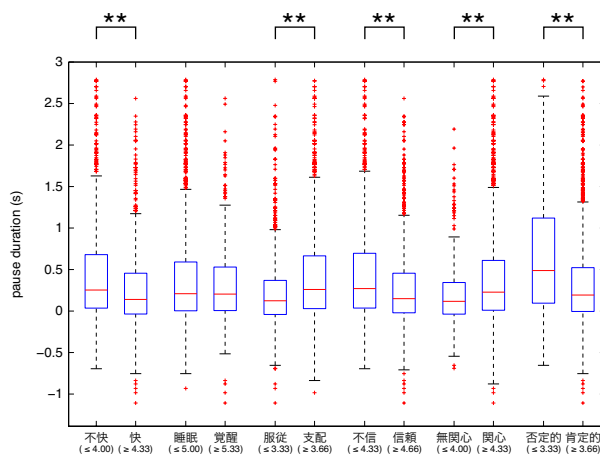


Fig. 3 Pause duration distribution of two groups split by each abstract dimension. (\*\*:  $p < 0.01$ )

### 3 パラ言語情報の効果

当該発話のパラ言語情報と発話間ポーズ長との関係を検討するため、6抽象次元のそれぞれに対し、ラベラ3名の平均評価値の大小で全データを2分割した。分割点は、分割された2群のデータ数に極端な偏りがない範囲でF値が極大となる点とした。各群の発話間ポーズ長の分布をFig. 3に示す。6抽象次元のうち、快-不快、支配-服従、信頼-不信、関心-無関心、肯定的-否定的の各次元の主効果が有意であった。すなわち、快・服従・信頼・無関心・肯定的発話の始端は先行発話の終端に対して早い傾向が、逆に不快・支配・不信・関心・否定的発話の始端は遅い傾向があった。

上の結果は、快-不快、信頼-不信、肯定的-否定的の各次元については事前の予想とよく一致するが、その他の次元についてはやや直感に反する。この原因の1つは、相槌様の発話「うん」だと考えられた。相槌は正常話者交替よりも早いタイミングで打たれる。先行発話断片からの平均ポーズ長は、正常話者交替では361 msであるのに対し、相槌では55 msである。一方、UUDBの中には聴覚的には相槌と変わらないがスラッシュ単位の途中以外に現れたために相槌とされていない発話が多く存在する。そこで、正常話者交替と分類された発話から「うん」だけからなるものを除いた1490発話に対して同様の分析を行った。その結果をFig. 4に示す。この結果から、「うん」を除く発話の始端は、快・覚醒・支配・信頼・関心・肯定的の方が早く、不快・睡眠・服従・不信・無関心・否定的の方が遅いことがわ

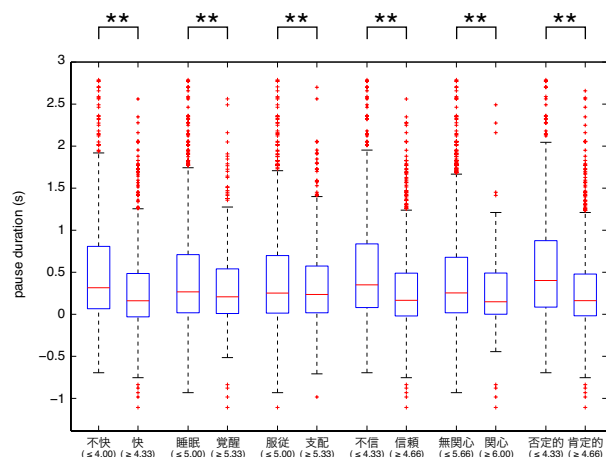


Fig. 4 Pause duration distribution of two groups split by each abstract dimension (excluding sole “uN”s). (\*\*:  $p < 0.01$ )

かった。

### 4 おわりに

自然な対話での話者交替における発話間ポーズ長を、UUDBを対象に分析した。その結果、発話間ポーズ長は、快・信頼・肯定的と知覚される音声では長く、不快・不信・否定的と知覚される音声では短い傾向があることがわかった。

UUDBに含まれる感情状態ラベリングでは発話間ポーズが呈示されていないため、パラ言語情報の知覚に発話間ポーズは関与していない。それにも関わらず上記の傾向が見られたことは、パラ言語情報の(知覚でなく)生成の段階で、発話タイミングと発話そのものの音響的特徴が、話者の感情状態や態度などから一貫した影響を受けていることを意味する。

今後は、発話タイミングの感情状態推定への応用可能性を検討したい。

### 参考文献

- [1] Koiso *et al.*, *Lang. & Speech*, 41(3-4), 295-321, 1998.
- [2] 岡登洋平 他, *情報処理学会論文誌*, 40(2), 469-478, 1999.
- [3] 大須賀智子 他, *人工知能学会論文誌*, 21(1), 1-8, 2006.
- [4] 長岡千賀 他, *人間工学*, 38(6), 316-323, 2002.
- [5] 石崎雅人, 伝康晴, “談話と対話,” 東京大学出版会, 2001.
- [6] 森大毅 他, *音講論 (秋)*, 311-312, 2007.