

日本語日常会話コーパスを用いた会話場面と声の高さの関係性の検討*
○石本祐一, △小磯花絵 (国語研)

1 はじめに

自発音声ではパラ言語情報や感情の影響によりピッチが様々に変動することが知られている [1] が, 日常生活の多様な状況を反映した音声データの不足により, 自発音声のピッチの多様性について大規模な定量的分析を行うことが困難であった。

国立国語研究所では, 2016 年より多様な種類の日常会話をバランス良く収録した大規模な日常会話コーパスとして『日本語日常会話コーパス』 (Corpus of Everyday Japanese Conversation, CEJC) の構築を進めている [2, 3]。本稿では, CEJC に収録される自発音声データを基に様々な会話場面における声の高さの違いを調べた結果について報告する。

2 日常会話データ

2.1 コーパス

2022 年の CEJC の本公開に先立ち, 2018 年 12 月より 50 時間分のデータがモニター公開された。Table 1 に, モニター公開データの対象となったインフォーマント (調査協力者) の属性を示す。40 代と 60 代の人数に 1 名の増減はあるものの, インフォーマントの年齢や性別が偏らないよう考慮して選定している。このようなインフォーマント達により日常生活の多種多様な場面で収録が行われた結果, 飲食店での友人との会話や車で移動中の家族との会話, 職場での同僚との会話など, これまでに観察することが困難であった幅広い日常会話が含まれることとなり, 50 時間のモニター公開データであっても自発会話音声の多様性を捉えることができると考えられる。

2.2 基本周波数推定

声の高さの分析には基本周波数 (F0) 情報が必要となるが, CEJC モニター公開版には F0 情報は付与されていない。そこで, STRAIGHT の音源情報分析関数 [4] を用い, モニター公開版の音声データから 1 ms 間隔で F0 抽出を行った。なお, 有声/無声判定は STRAIGHT から得られる非周期性情報を利用した。さらに, 推定エラーの影響を低減させるために, 推定された F0 のうち上位 10% と下位 10% の値を取り除いた。最終的に, 性差と個人差を正規化するため各話者ごとの平均値と標準偏差を用いて対数 F0 を Z-score に変換し, 発話単位の平均 F0 を求めた。

3 会話場面と声の高さ

インフォーマントの家族や友人等はそれぞれの会話場面に応じて登場するが, インフォーマントは必ず

Table 1: Attributes of informants. The informants were almost balanced in terms of sex and age.

話者 ID	年齢	性別	職業	
K003	20-24	女性	大学生	*
T006	25-29	男性	大学院生	*
T009	20-24	女性	大学生	*
T010	20-24	男性	大学生	*
K001	35-39	女性	会社員・公務員等	*
T001	35-39	男性	自営業・自由業	*
T003	35-39	女性	専業主婦	*
T005	35-39	男性	会社員・公務員等	
C001	40-44	女性	会社員・公務員等	*
K004	40-44	女性	パート・アルバイト	*
T002	40-44	男性	自営業・自由業	*
T011	40-44	女性	パート・アルバイト	
T016	40-44	男性	会社員・公務員等	
C002	55-59	女性	会社員・公務員等	*
K002	50-54	女性	自営業・自由業	*
S001	50-54	男性	会社員・公務員等	*
T015	50-54	男性	会社員・公務員等	*
T004	60-64	女性	専業主婦	*
T007	70-74	男性	定年退職	*
T013	65-69	男性	その他 (非常勤講師)	*

すべての会話場面に参与している。そこで, 本研究の分析ではインフォーマントの音声のみを対象とする。

3.1 発話の向け先との関係による違い

まず, 日常会話において, 話者とその発話の向け先の間の社会的な関係によって F0 が異なるかどうかを調べた。ここで, 発話の向け先の種別として, 子ども・配偶者・父母・兄弟姉妹・友人・同僚・取引先・客の 8 種類を取り上げることとした。また, 向け先の種別の同定を簡便にするために, 同種の関係となる参与者だけが存在する会話に限定して分析を行った。例えば, インフォーマント (話者) が父親である場面において他に子どもだけが参与している会話場面は分析対象となるが, 妻と子どもが参与している会話場面は分析対象外となる。その結果, Table 1 の行末に * を付した 17 名のインフォーマントによる発話が分析対象となった。

Fig. 1 に, 話者の F0 と発話の向け先の種別との関係を示す。TukeyHSD による多重比較の結果, 次のような傾向を持つグループに分類された。(1) 子ども・配偶者・父母に対して F0 は低くなる。(2) 友人に対しての発話がもっとも高くなる。(3) 同僚・取引先・客に対しては, 友人よりは低く近親者よりは高くなる。すなわち, 日常会話の声の高さは兄弟姉妹を除い

* An investigation of relationships between fundamental frequencies and conversation situations based on the corpus of everyday Japanese conversation. by ISHIMOTO, Yuichi and KOISO, Hanae (National Institute for Japanese Language and Linguistics)

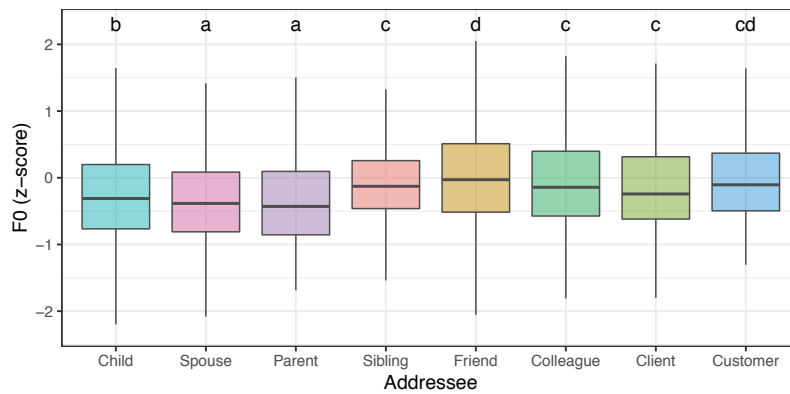
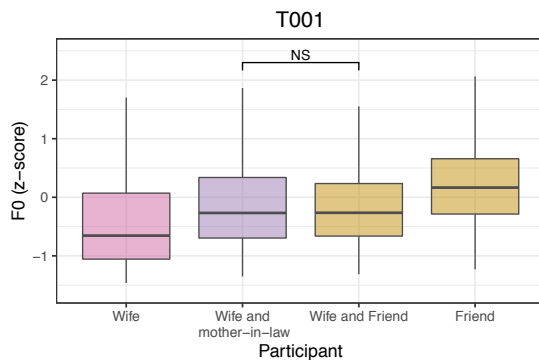
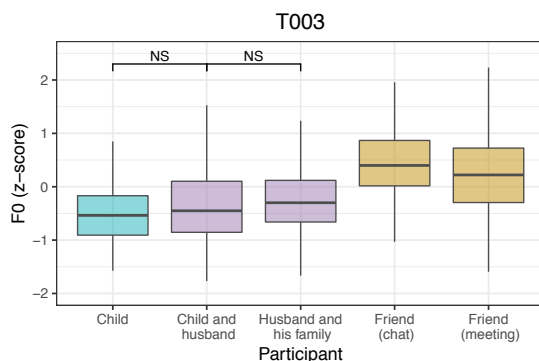


Fig. 1: Distributions of the average F0 of utterance for relationships between speaker and addressee. The letters above the plots indicate that groups with the same letter are not a detectable difference in the Tukey HSD test ($p < 0.05$).



(a) Informant T001



(b) Informant T003

Fig. 2: Distributions of the mean F0 for types of participant.

た近親者に対しては低く、それ以外には高くなる傾向がある。

3.2 会話場面の同席者による違い

次に、異なる種別の参加者が存在する会話場面においても声の高さの変化に同様の傾向が見られるのかを、インフォーマントごとに調べた。本稿では30代のインフォーマント T001 と T003 を取り上げる。

Fig. 2(a) にインフォーマント T001 の発話の F0 とそれぞれの会話場面の参加者（同席者）との関係を示す。妻と義母がいる会話場面と妻と友人がいる会話場面の間では F0 に有意差が見られなかったが、妻だけの場面は低く、友人だけの場面では高い発話となっていた。すなわち、妻だけの場合と比べて同居して

いない義母や友人が加わった状況で高くなり、妻がいない友人だけの状況でもっとも高くなることから、同席者への心理的な距離が声の高さに影響しているように思われる。

Fig. 2(b) にインフォーマント T003 の発話の F0 とそれぞれの会話場面の参加者との関係を示す。T001 と同様に子どもや夫といった家族だけが同席している場合は低く、友人に対しては高くなる。さらに、友人に対する発話でも雑談の場合と会合の場合では高さが異なり、同席者の種別だけではなく状況によっても違いが現れることがわかる。また、T001 とは異なり、夫の家族が同席していても子どもや夫に対する場合とははっきりした違いが見られなかった。これは、T001 の義母に対する心理的距離と T003 の夫の家族に対する心理的距離による違いとも考えられるが、より詳細な分析が必要である。

4 おわりに

様々な状況で収録された日常会話音声に基づき、発話の向け先や同席者との関係による声の高さの違いを調べた。その結果、子どもや配偶者、父母といった家族に対しては低く、取引先や客など丁寧さが必要な相手や友人には高い声で話していることが示された。また、発話の直接の向け先だけではなく、会話場面に同席している参加者の属性によっても声の高さが変わることが観察された。

謝辞 本研究は、国立国語研究所「大規模日常会話コーパスに基づく話し言葉の多角的研究」プロジェクト、および JSPS 科研費 16H03421, 17H00914, 18K11514, 19H01252 の助成を受けたものです。

参考文献

- [1] 森 他 “音声は何を伝えているか,” コロナ社, 2014.
- [2] Koiso *et al.*, Proc. LREC2018, 4259–4264, 2018.
- [3] 石本 他, 音講論 (秋), 811–812, 2018.
- [4] https://github.com/HidekiKawahara/YANGstraight_source