

日本語日常会話コーパスから見える会話場面と声の高さの関係性

著者	石本 祐一
雑誌名	言語資源活用ワークショップ発表論文集
巻	5
ページ	371-378
発行年	2020
URL	http://doi.org/10.15084/00003177

日本語日常会話コーパスから見える会話場面と声の高さの関係性

石本 祐一 (国立国語研究所コーパス開発センター) *

Relationships between fundamental frequencies and conversation situations based on the corpus of everyday Japanese conversation

Yuichi Ishimoto (National Institute for Japanese Language and Linguistics)

要旨

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』(CEJC)の構築を進めている。CEJCに収録される自発音声のうち50時間の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ところ、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。

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(森ほか 2014)が、日常生活の多様な状況を網羅するような音声データはこれまでなく、日常の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が行われることはなかった。そのような分析に活用できる日常会話データを補うために、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である『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(Koiso et al. 2018, 小磯ほか 2019)。本稿では、日常会話における韻律の多様性を示すことを目的として、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

CEJCは日常生活で自然に生じる会話を収録することが念頭に置かれており、研究者の介入による影響を排除するために、インフォーマント(研究協力者)自身による映像の録画および

* yishi@ninjal.ac.jp



図1 コーパスに収録されている映像データの例

音声の録音が行われている。図1にコーパスに収録されているデータの一場面を示す。360°カメラを中央に置き、対角にGoProカメラを配置して録画することで、会話の参加者の様子が様々な角度から観察できるように考慮している。また、音声の録音については、日常生活の自然な振る舞いを妨げるヘッドセットマイク等の利用は避け、各参加者がICレコーダを胸の位置に配置して、それぞれの音声を話者になるべく近いマイクで録音することで会話音声の重なりや周囲の雑音の入り込みを防ぐことを企図している。しかし、収録状況や参加者間の位置関係によって、他の話者の音声や雑音が入り込んでしまう場合もあり、常にクリーンな音声を得られているわけではない。

2022年のCEJCの本公開に先立ち、2018年12月より50時間分がモニター公開されている。表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されている。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のある程度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、音声分析合成システムSTRAIGHTの音源情報分析関数(Kawahara et al. 2017)を用い、モニター公開版の音声データから1ms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

表1 インフォーマントの属性

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）	*

数 F0 を Z-score に変換し、発話単位の平均 F0 を求めた。なお、発話単位として、コーパスに付与されている「長い発話単位」を用いている。

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すべての会話場面に参与している。そこで、本稿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と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加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、表 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。

図 2 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

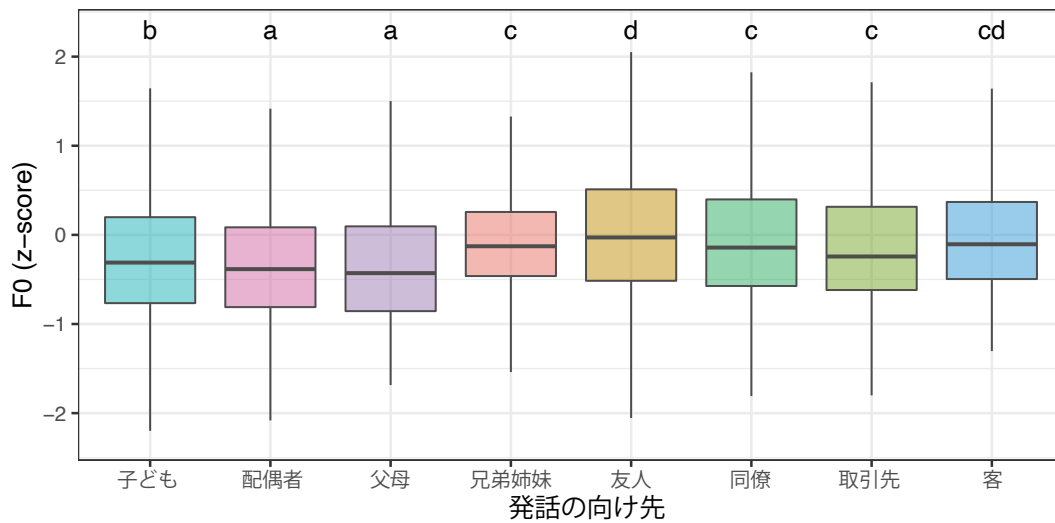


図2 発話の向け先ごとの発話の平均 F0 の分布。グラフの上に付した文字が TukeyHSD 検定 ($p < 0.05$) によるグループを表しており、同じ文字であれば有意差がないことを意味する

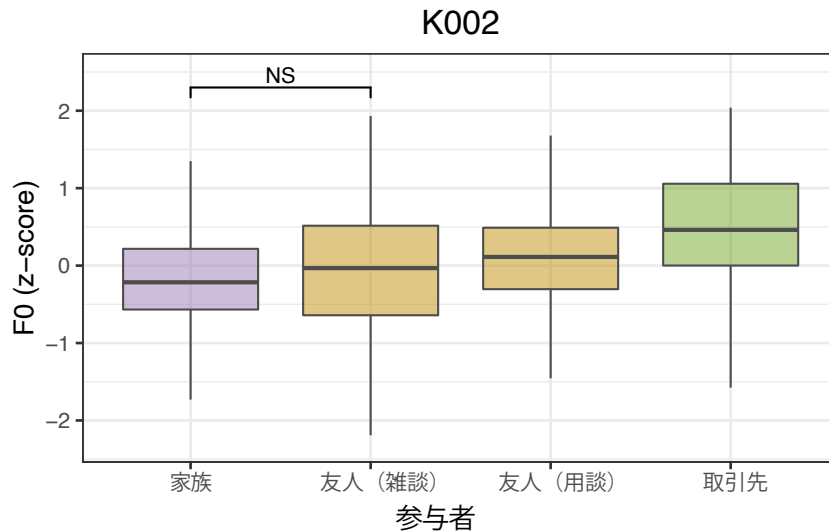


図3 インフォーマント K002 の F0 と会話場面の参加者との関係

果、次のような傾向を持つグループに分類された。

- 子ども・配偶者・父母に対して F0 は低くなる
- 友人に対しての発話をもっとも高くなる
- 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た近親者に対しては低く、丁寧さが必要となる同僚・取引先・客相手には高くなり、友人に対してはさらに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、5人のインフォーマントの結果を

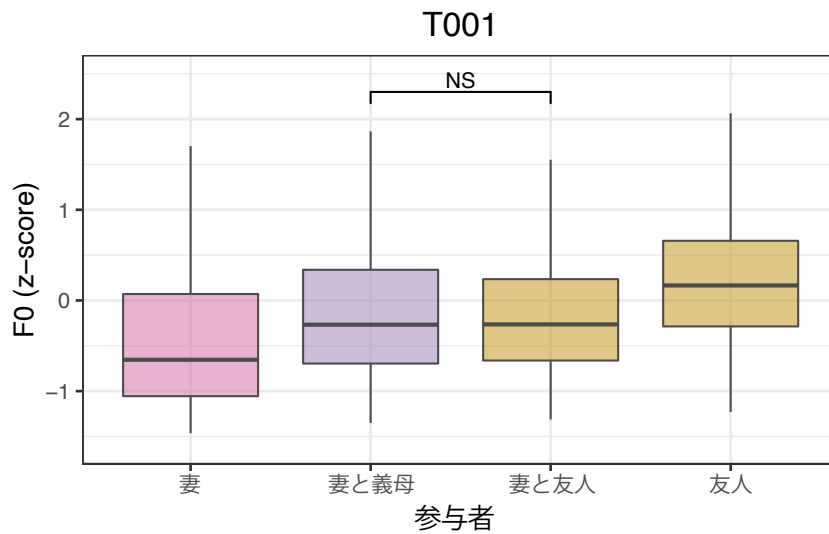


図4 インフォーマント T001 の F0 と会話場面の参加者との関係

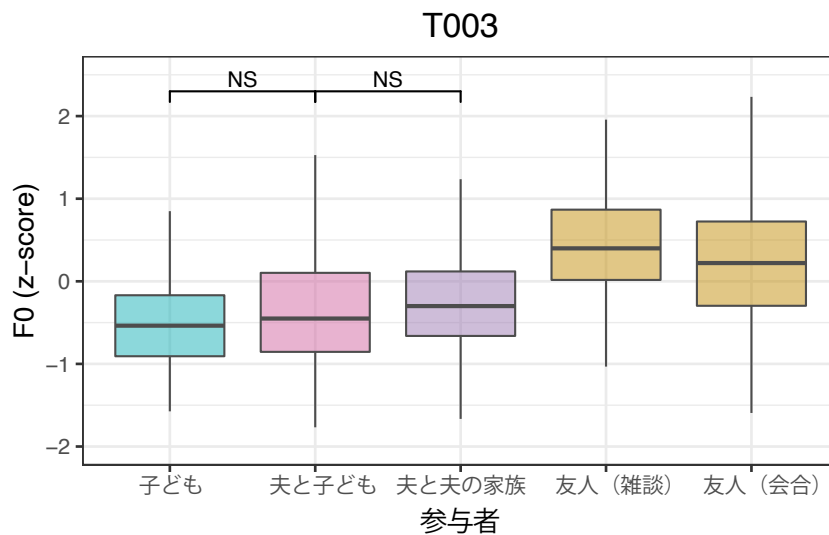


図5 インフォーマント T003 の F0 と会話場面の参加者との関係

取り上げる。

図3にインフォーマント K002の発話のF0とそれぞれの会話場面における参加者(同席者)との関係を示す。家族との会話場面と友人との雑談場面の間ではF0に有意差が見られないが、友人との打ち合わせや取引先との会合ではF0が高くなっている。特に取引先との会合が一番高くなっており、丁寧さが必要な場面でF0がより高くなる傾向にあるといえる。一方、友人との雑談でF0が低くなっていることから、3.1節で示した全体的な傾向とは異なる結果が見られた。

図4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における参加者との関係

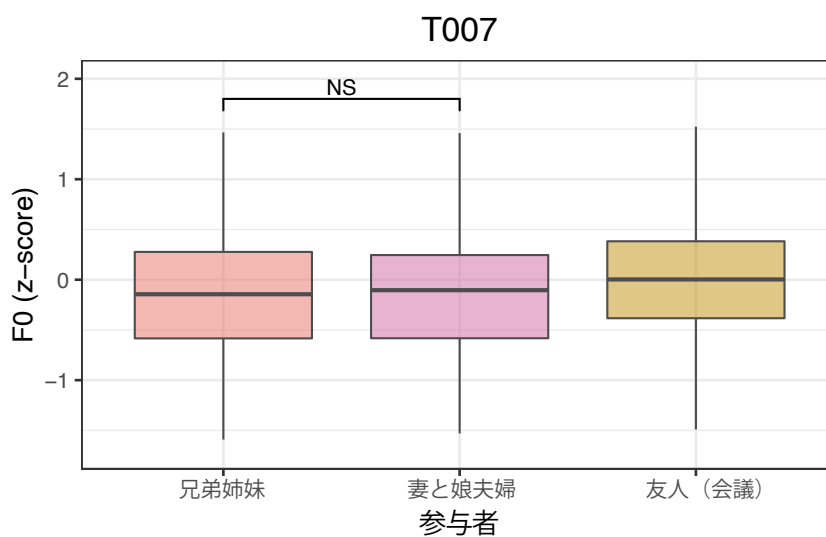


図6 インフォーマント T007 の F0 と会話場面の参与者との関係

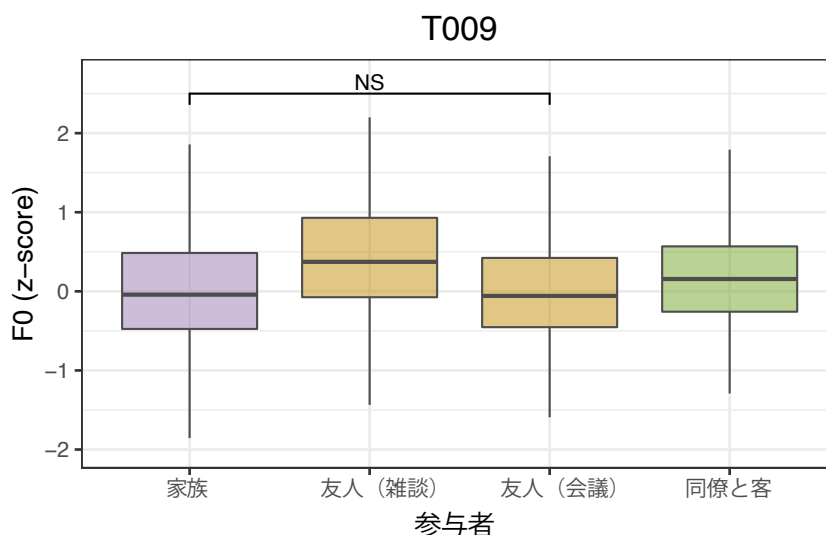


図7 インフォーマント T009 の F0 と会話場面の参与者との関係

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。

図5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における参与者との関係を示す。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が現れている可能性がある。

図6にインフォーマントT007の発話のF0とそれぞれの会話場面における参与者との関係を示す。兄弟での会話場面と妻や娘夫婦との会話場面においてF0に有意差は見られず、友人との会話ではわずかながらF0が高くなっている。つまり、この話者に関しては、参与者との関係によってF0が顕著には変わらない傾向にある。

最後に、図7にインフォーマントT009の発話のF0とそれぞれの会話場面における参与者との関係を示す。この話者は友人との雑談と同僚・客との会話でF0が高くなっており、3.1節に示した傾向に近い。しかし、友人との打ち合わせではF0が高くなっておらず、丁寧さが必要な場面で高くなるという全体傾向とは反対の結果となった。

3.3 議論

3.1節では、日常会話における発話のF0は、子どもや配偶者などの近親者ほど低くなること が示された。一方、家族以外への発話では、F0が高くなる傾向があった。これは、社会的規範としての相手への礼儀正しさを高いF0によって無意識に表現しているからではないだろうか。また、家族と友人は親しみやすさという点で類似しているが、F0に関しては反対の傾向を示しており、単純な親近感という要因では説明できない。

3.2節では、発話の向け先だけでなく、他の参与者も発話のF0に影響を与えることが示された。例えば、図4に示すように、インフォーマントT001は妻だけがいる状況ではF0が低い のに対し、義母もいる状況ではF0が高い。義母は家族の一員ではあるが実際には自分と妻とは同居していないため、義母を同居家族とは区別して扱っていることを示唆している。また、友人とだけの場合にはF0が高く、妻と友人の両者がいる場合にはF0が低くなっており、妻の存在によってF0が調整されている可能性がある。一方、インフォーマントT003とT009は、友人との会話場面であっても雑談と打ち合わせでは異なるF0で発声していた。また、インフォーマントK002とT009で見られた接客に着目すると、友人との会話よりも取引先や客との会話の方がF0が高いことがわかった。

以上をまとめると、インフォーマントや参与者によってF0の変化の度合いは異なるが、次の点が共通して見られた。

- 家族に対しては、比較的低いF0で話す
- 顧客や取引先などの場面では、高いF0で話す

本分析により、実生活での発話のF0の違いは参与者や会話の状況との複雑な関係性から生じるものであり、日常会話では発話の韻律的な多様性が顕著に現れることが示された。もちろん、会話におけるF0の違いを説明する要因は他にもあるため、今後は会話が行われる場所や詳細な話題も考慮した分析を行う必要がある。

4. おわりに

日本語日常会話コーパスを用いて、実生活における自発発話の韻律的多様性を観察するために、話し相手の属性に着目して声の高さに関わる音響特徴量である F0 を分析した。その結果、家族に向けた発話は一般的に低い F0 で発声され、丁寧な発話ほど F0 が高くなることがわかった。また、会話場面に同席している参加者を考慮して F0 を分析したところ、参加者の影響による F0 の変化割合は個人差があり、参加者との複雑な関係から韻律的多様性が生じていることが示された。

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 17H00914, 18K11514 の助成を受けたものです。

文 献

- 森大毅・前川喜久雄・粕谷英樹 (2014). 『音声は何を伝えているか：感情・パラ言語情報・個人性の音声科学』音響サイエンスシリーズ/日本音響学会編:12 コロナ社.
- Hanae Koiso, Yasuharu Den, Yuriko Iseki, Wakako Kashino, Yoshiko Kawabata, Ken'ya Nishikawa, Yayoi Tanaka, and Yasuyuki Usuda (2018). "Construction of the Corpus of Everyday Japanese Conversation: An Interim Report." *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, pp. 4259–4264.: European Language Resources Association (ELRA).
- 小磯花絵・天谷晴香・石本祐一・居關友里子・白田泰如・柏野和佳子・川端良子・田中弥生・伝康晴・西川賢哉 (2019). 「『日本語日常会話コーパス』モニター公開版の設計と特徴」言語処理学会第 25 回年次大会発表論文集, pp. 367–370.
- Hideki Kawahara, Ken-Ichi Sakakibara, Masanori Morise, Hideki Banno, and Tomoki Toda (2017). "Accurate estimation of f0 and aperiodicity based on periodicity detector residuals and deviations of phase derivatives." *2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pp. 1556–1564.

関連 URL

『日本語日常会話コーパス』

<https://www2.ninjal.ac.jp/conversation/cejc.html>

音声分析合成システム STRAIGHT

https://github.com/HidekiKawahara/YANGstraight_source