

2021 年度
プレ卒論

非流暢性を通じた対話場面における
パラ言語情報の表出に関する考察

指導教員: 菊池英明

早稲田大学 人間科学部

学籍番号:1J19E058-1

加藤隆聖
Ryusei Kato

目次

1	はじめに	3
1.1	背景	3
1.2	目的	6
1.3	波及効果	6
2	予備的なプロセス	7
2.1	予備的なプロセスについて	7
2.2	使用データ	7
2.3	非流暢性のアノテーション	7
2.4	結果	7
2.5	考察	8
3	まとめ	9
3.1	おわりに	9
	参考文献	10

図目次

1	音声による情報伝達過程の模式図	3
2	パラ言語情報を考慮した音声生成過程の模式図	4
3	The speech chain	4
4	発話向け先ごとの発話の平均 F_0 の分布	5
5	対話に出現する非流暢性の半年前と半年後の出現頻度	8

1 はじめに

1.1 背景

対話場面における情報伝達は複雑である。これは音声には様々な情報が含まれており、会話時には話し手と聴き手の間に高度なやりとりがなされているためである。音声に含まれる情報は、言語情報、パラ言語情報、非言語情報の三つに分類することができる。以下に藤崎による3分類を引用する。言語情報とは、書き言葉によって明示的に表現される情報や文脈から一意かつ容易に推測することが可能である情報のことを指す。パラ言語情報とは、書き言葉に転写すると推測不可能となる情報で、言語情報を補助ないし変容するために話者が意図的に生成する情報のことを指す。特に発話に込められた話者の意図、態度や発話のスタイルなどが該当する。非言語情報とは、話者の年齢、性別、個人性、身体ないし感情の状態などの要因に関わる情報のことを指す。これらの要因は、発話の言語的・パラ言語的内容とは直接に関係せず、話者が意図的に制御することも一般には不可能である [1]。これら三つの情報を伝達するときに、話し手と聴き手の音声伝達の模式図を図1に示す。図1では、聴き手においてパラ言語情報と非言語情報が必ずしも正確に知覚されるとは限らないため、破線の矢印で示されている。パラ言語情報の影響は発話全体にわたって顕在化することもあるが、むしろ多くは発話の冒頭や末尾において局所的に顕在化している [2]。パラ言語情報を分析するためには、発話全体を対象とすることが有効であることが示されている。

このように、音声伝える情報について研究が行われてきたが、パラ言語情報や非言語情報がどのような仕組みで生成され、伝達されているのかについて全てを明らかにすることはできていない。

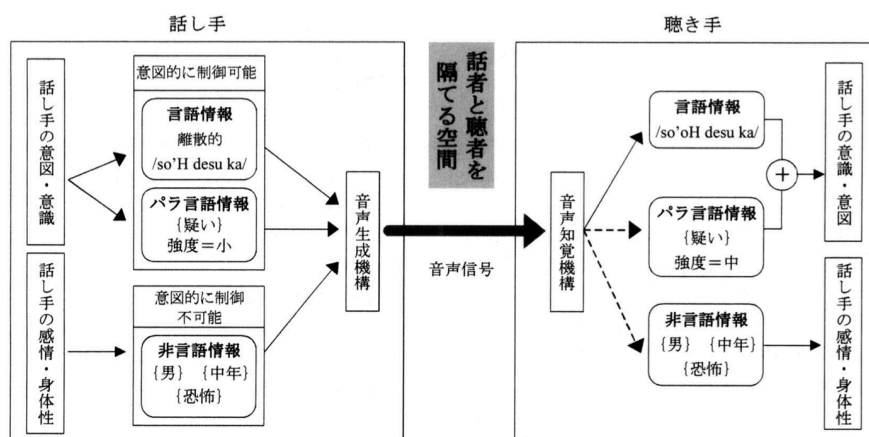


図1: 音声による情報伝達過程の模式図。[4]より抜粋。話し手が聞き手へ言葉発した時の情報の伝達方法を模式的に表している。話し手発した情報の中でパラ言語情報と非言語情報が必ずしも正確には伝わらないことが示されている。

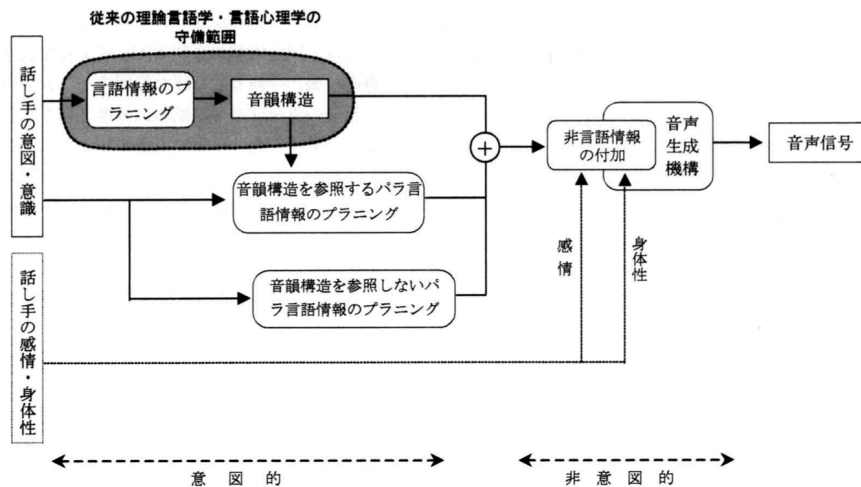


図 2: パラ言語情報を考慮した音声生成過程の模式図。[2] より抜粋。話し手が音声を発する時に、言語情報、パラ言語情報、非言語情報がどのように組み合わせて生成されているのかについて模式的に表されている。この画像から感情と身体性を含む非言語情報の付加は非意図的に生成されていることが示されている。

パラ言語情報を考慮した音声生成過程の模式図を図 2 に示す。図 2 によると、発話する際の音声に含まれる情報には意図的なものと非意図的なものが存在する。さらに、意図的なものに含まれるパラ言語情報には音韻構造を参照するものと参照しないものに分けられている。このように音声のみに対してアプローチしている研究はなされているが、発話者同士の関係や発話者のスタイルに着目はされていない。

話した言葉が話し手の脳から聞き手の脳へ伝わる過程の模式図を図 3 に示す。図 3 では、話し手から発せられた音声と聴き手が音声を知覚するまでの過程がつながりを持っていることを示している。よって、話し手の意図と聴き手が知覚した内容には齟齬が生じることがある。ゆえに、パラ言語情報を認知すること自体やパラ言語情報の表現の仕方が曖昧なのである。

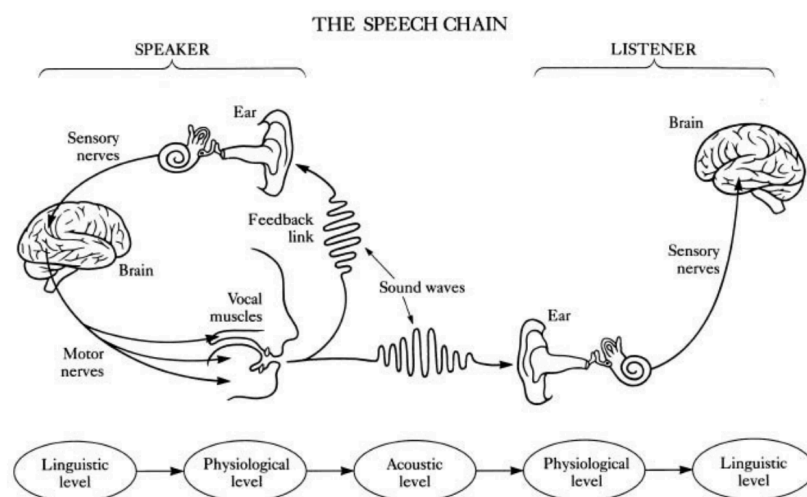


図 3: The speech chain。[13] より抜粋。話した言葉が話し手の脳から聞き手の脳へ伝わる過程が表されている。

Wertsch によると、自分の音声は他者の声から切り離されて存在するのではなく、発話が向けられる他者の声に基づいて作られるのである。また、音声は他者との関係において成り立つのであり、音声には状況依存性がある [3]。その状況とは、話し手と聴き手が物理的環境を共有することである [4]。

前川らによると、対話音声には、発話の重複・あいづち・語形の縮約・言いさし・言い直し・繰り返しなどの特徴的現象が観察される。これらの多くは対話が複数の参加者による協働行為であることに起因している [5]。さらに、石本によると、発話向け先との関係の違いによって日常対話の声の高さが変化することがわかっている。また、会話場面の同席者している参加者との関係によって韻律的多様性が生じていることがわかっている [6]。発話向け先との関係の違いや同席している図を 4 に示す。図 4 によると、家族に向けた発話は一般的に低い Fo で発声され、丁寧な発話ほど Fo が高くなることが読み取れる。つまり、対話相手によって発話スタイルが変わるため、パラ言語情報の表出の仕方が変わるのである。よって、パラ言語情報が発話者のみによって発せられるのではなく、発話相手によってパラ言語情報の表出の仕方が変わるため、音声のみに注目してはいけなくと考えられる。音声のみに注目するのではなく、対話者の発話のスタイルや対話場面についても考慮すべきであると考えられる。

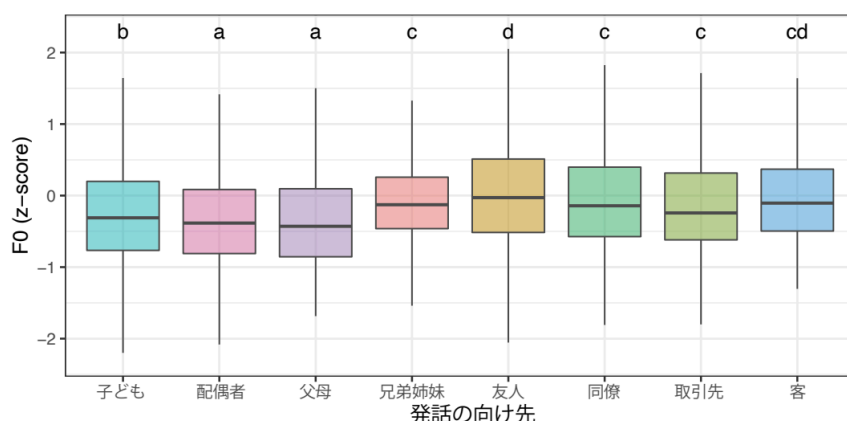


図 4: 発話向け先ごとの発話の平均 Fo の分布。 [6] より抜粋。

音響特徴量は感情ラベルに比べコーパスの特性の違いの影響を受け易いことがわかっている [7]。よって、従来の対話音声研究では、パラ言語情報が収録したコーパスの特性の影響を受けやすいため、パラ言語情報を普遍的に明らかにするには至らなかった。しかし、国立国語研究所では 2016 年より、日常生活の中で自然に生じる多様な場面・話者の会話 200 時間を対象とする『日本語日常会話コーパス』(CEJC) の構築が進んでいる [8]。多様な関係性における対話を収録した日常対話コーパスによって、コーパスの特性に限定されことなく対話音声を分析できるようになった。

次に、自発発話について述べる。自発発話には非流暢性が必ず伴う。よって、自発発話の分析には非流暢性の分析が必要であると考えられる。また、非流暢性は自発発話に現れるという特徴だけでなく、コミュニケーションにおいて様々な役割がある。先行研究では以下のように述べられている。聞き手に対して様々な影響を与える現象である。特に、フィラーは発話権を保持する役割を持っているおり、音韻の引き伸ばしは発話の中断の合図になっている [9]。このように非流暢性はコミュニケーション方略として用いられていることはわかっている。しかし、非流暢性の全体像をカバーするような分析は行われていない。さらに、言語情報とパラ言語情報との間の交互作用について分析するべきであるということは、[2] でも述べられている。

1.2 目的

本研究では、日常的かつ自然な対話音声と非流暢性の関係を明らかにすることを目指す。さらに、様々な関係性や発話のスタイルによってパラ言語情報の表出の仕方がどのように変化するのかを調べる。この目的を達成するために、次の三つのリサーチクエスチョンを設定した。

1. 親密度の変化によって非流暢性はどのように変化するのか
2. 非流暢性の出現頻度にはどのような傾向があるのか
3. 対話において非流暢性が果たす役割の定量的な特徴はどのようなものか

1.3 波及効果

パラ言語情報を構成する要素を明らかにすることで、パラ言語情報を分析する際の一つの指標として確立することができる。また、パラ言語情報の表出方法の仕組みを明らかにすることにもつながるであろう。さらに、スタイルの違いに応じたパラ言語情報の表出方法がどのようなものであるかを明らかにできる。その結果、場面に適した音声を使うことで、教育やトレーニングを行う際にも、どのような音声が指導に適切であるか知ることができる。それに加え、パラ言語情報についての知見を応用することで、音声合成、音声変換の多様性を実現できるであろう。

2 予備的なプロセス

本章では、本研究での予備的なプロセスを述べる。

2.1 予備的なプロセスについて

そもその問題として日常対話コーパスを扱うときに、対話場面の違いをどのように体系化することができるのかという問題がある。日常の中の対話場面での関係の変化として親密度の変化が生じることが考えられる。先行研究の中で、音声に付随する書き起こしテキストに対して、speaking style の自動推定のために、明瞭さ・親しさ・社会階層の 3 尺度を用いて推定を行なっている [10]。書き言葉に対して親密度の変化が重要であるのと同様に、話し言葉においても親密度の変化は重要な要素であるといえる。親密度が異なる場面で非流暢性の特徴がどのように変化するかを調べる。非流暢性の特徴を明らかにすることで、対話場面の違いを非流暢性の特徴から分類することができると考えられる。また、非流暢性によって場面の違いを数値化することで、多種多様な場面を非流暢性という観点から比較検討することができるようにとなると考えられる。

2.2 使用データ

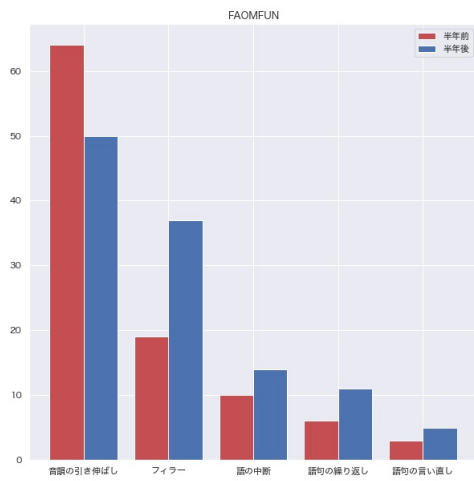
使用データは、同研究室で収集された『新入生対話コーパス』(FDC)[11] の 2011 年の対話を使用する。本コーパスは、比較的初対面である大学生二名が行った対話と半年後に同じ二名で行った対話が収録されている。2011 年の対話を今回の分析に使用した理由は、1 度目の対話が完全に初対面であるためである。分析に用いたデータは、FDC から男性と男性の対話 2 組、女性と男性の対話 2 組、合計 8 対話である。FAOMFU と FAOMNA は男性対女性の対話であり、MAKMTK と MTKMTT は男性対男性の対話である。

2.3 非流暢性のアノテーション

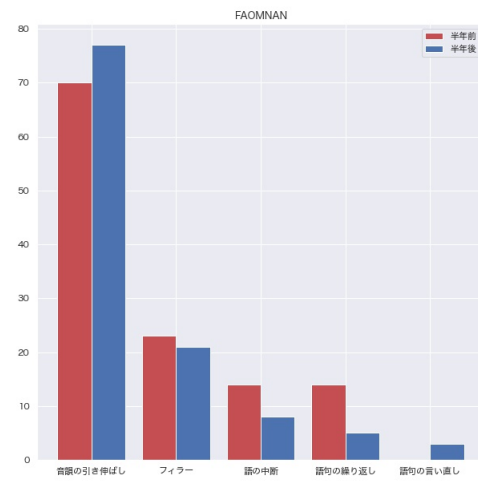
アノテーションには、伝らの非流暢性の分類を使用した。以下にその分類を示す。音韻の引き伸ばしとは、発話の切れ目において母音を持続して発話することである。フィラーとは、発話における無意味語のことである。語の中断とは、ある単語の途中で言葉を言い止めることである。語句の繰り返しは、同じ単語を複数回繰り返すことである。語句の言い直しとは、一度発話した単語を違う単語に言い直すことである [9]。具体例を以下に示す。「なんか、その、俺の相手した人がー、俺よりふたまわりくらい大きいー、なんか、ちゅうがっ、中学の頃ー、その県で、ゆっ柔道で優勝した人だったの、で、おれが間違えてそいつなげちゃってー」[11] 音韻の引き伸ばしとは「人がー」「大きいー」にあたる。フィラーとは、「その」にあたる。語の中断とは「ちゅうがっ」にあたる。語句の言い直しとは「ゆっ」にあたる。なお、フィラーの分類は『日本語話し言葉コーパス』(CSJ) の書き起こし基準に準ずる [12]。

2.4 結果

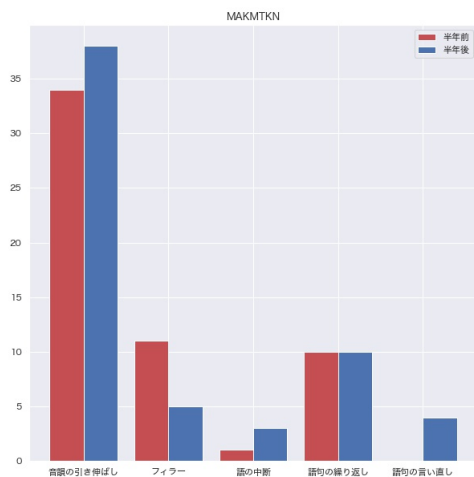
対話に現れる非流暢性の出現頻度を図 5 に示す。男性対女性の対話の半年前と半年後の非流暢性の出現頻度を図 5a、図 5b に、男性対男性の対話の半年前と半年後の非流暢性の出現頻度を図 5c、図 5d に示す。



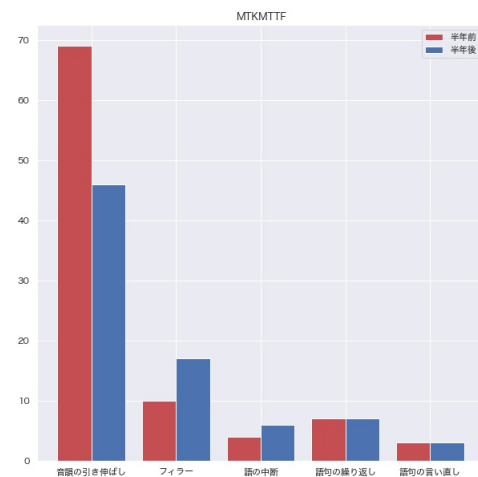
(a) FAOMFU



(b) FAOMNA



(c) MAKMTK



(d) MTKMTT

図 5: 対話に出現する非流暢性の半年前と半年後の出現頻度

2.5 考察

半年前と半年後の親密度の変化によって、非流暢性が一様に変化するという結果は得られなかった。しかし、非流暢性の出現頻度は個人差の影響が大きいことから、個人に着目することでどのような特徴があるのかについて明らかにできると考える。また、男性対女性の対話と男性対男性の対話に分けて分析することで、非流暢性の変化の傾向を知ることができると考える。

3 まとめ

本章では、本研究のまとめを述べる。

3.1 おわりに

全体の傾向として、親密度が高まることで非流暢性が変化するという傾向は得られなかったが、個人ごとの変化を見ることで何らかの傾向が得られると考えられる。今後は個人ごとの非流暢性の変化を調べ、さらに非流暢性と親密さの関係を音響特徴量から明らかにしていく。

参考文献

- [1] Fujisaki, Hiroya. "Prosody, models, and spontaneous speech." *Computing prosody*. Springer, New York, NY, 1997. 27-42.
- [2] 前川喜久雄, and 北川智利. "音声はパラ言語情報をいかに伝えるか." *認知科学* 9.1 (2002): 46-66. 17.1: 109-118.
- [3] Wertsch, James V. *Voices of the mind: Sociocultural approach to mediated action*. Harvard University Press, 1993. pp. 46 - 63 (ジェームス・ワーチ著, 田島信元ほか訳 (2004) 心の声, 福村出版, 東京, pp. 67 - 89)
- [4] 前川喜久雄 (1998). 音声学. 『岩波岩波講座言語の科学 第2巻, 音声』, 1-52, 岩波書店.
- [5] 前川喜久雄. "韻律とコミュニケーション." *日本音響学会誌* 55.2 (1999): 119-125.
- [6] 石本祐一. "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." *言語資源活用ワークショップ発表論文集= Proceedings of Language Resources Workshop. No. 5*. 国立国語研究所, 2020.
- [7] 永岡篤, 森大毅, and 有本泰子. "感情音声コーパス共通化のための新たな感情ラベル推定における既存感情ラベル併用の効果." *日本音響学会誌* 73.11 (2017): 682-693.
- [8] 小磯花絵, 菊池英明, and 山田高明. "『日本語日常会話コーパス』への韻律ラベリング-ラベリングの設計と日常会話の韻律の特徴." *人工知能学会研究会資料 言語・音声理解と対話処理研究会*. 一般社団法人人工知能学会, 2020.
- [9] 伝康晴, and 渡辺美知子. "音声コミュニケーションにおける非流暢性の機能 (特集) 音声伝達する感性領域の情報の諸相)." *音声研究* 13.1 (2009): 53-64.
- [10] 沈睿, and 菊池英明. "音声言語コーパスにおける speaking style の自動推定—転記テキストに着目して—." *自然言語処理* 21.3 (2014): 445-464.
- [11] 新入生対話コーパス (FDC), kikuchiken-waseda, 2018
- [12] 小磯花絵, and 前川喜久雄. "『日本語話し言葉コーパス』の設計の概要と書き起こし基準について." *情報処理学会研究報告音声言語情報処理 (SLP) 2001.55 (2001-SLP-036) (2001): 1-8*.
- [13] Denes, Peter B., Peter Denes, and Elliot Pinson. *The speech chain*. Macmillan, 1993.