

沈黙の数理モデル、あるいは天使の通る確率について

PC で苔を育てる人

2025 年 4 月 18 日

概要

フランス語の“Un ange passe”（天使が通る）は、集団において不意に沈黙が訪れる現象を指す慣用句である。このような表現が世界中で見られることから分かるように、この現象は人間の会話の持つ普遍的な特徴の一つである。しかし一方で、その沈黙が持つ意味は、国や文化、あるいはコミュニティ内の関係性によって大きく変化する。そのため、組織の生産性に対する影響はもちろん、アイデア・創造性との関係や、あるいは医療現場における効果などが着目され、様々な研究が行われてきた。このような沈黙の持つ重要な役割にもかかわらず、数理モデルによってその特徴を明らかにしようとした研究はこれまでほとんど行われてこなかった。そこで本研究では、マルコフ連鎖モデルによって集団の沈黙を記述し、それを解析することを試みた。得られた結果から、各々が周囲の会話に対する意識を下げていくと、あるところで沈黙の発生確率が急激に 0 となる、いわゆる相転移のような現象が見られることが分かった。言い換えれば、このような沈黙は、皆がわずかにでも周囲を意識することで初めて発生する。ここで提案するモデルは、会話に対するより深い理解をもたらすだけでなく、異文化交流や、組織における生産性、あるいは医療現場などにも寄与できる可能性がある。

1 導入

フランス語の“Un ange passe”は、直訳すると「天使が通る」であり、これは会話が不意に途切れた際の静寂を意味する表現である。同様の言い回しは、ドイツ語の“Ein Engel geht durchs Zimmer”（天使が部屋を通り抜ける）、英語の“An angel is passing”などにも確認でき、古くは、古代ギリシャにおけるプルタルコスの言明、「集まりの中で突然静寂が訪れるとき、人々はヘルメスが部屋に入ってきたと言う」に見られる [1]。また、オランダ語では、“Er geht een dominee vorbei”（牧師が通り過ぎる）、ロシア語では“Вышел дурак”（愚か者が生まれた）が、それぞれ突然の静寂を表す慣用句として用いられることがある。このように、多くの異なる言語に同様の状況を指す言葉が存在するという事実は、集団に突如訪れる沈黙が、人類にとって普遍的な現象である可能性を示唆する。

沈黙は単に発言していないという以上に、何らかの意味を持つことがあり、その解釈・価値観は文化によって異なる [2–10]。例えば、ニューヨークのユダヤ人の会話スタイルを分析した研究では、彼らが沈黙を否定的に評価し、そのため会話の流れを維持し続ける傾向があることが示されている [2]。一方、西部アパッチ族の社会における沈黙は、関係性の不安定さを反映する [11]。これは、社会的な状況に依存し、例えば初対面の場合には、相手の社会的地位を尊重する行動として機能し、若者の付き合い始めにおいては、緊張や恥じらいの現れと認識される [11]。また、親子が再会する場面では不安や距離感を示し、争いごとにおいては自己防衛の手段となる [11]。そして、死者を悼むときは、沈黙は悲しみを共有する [11]。日本では、沈黙が単に拒否を意図する場合がある一方で、そこに根付く「雄弁は銀、沈黙は金」という言葉が示す通り、沈黙に対してポジティブな印象もある [6]。それは、対立を避け、調和を保つ手段として利用される [6]。アフリカの人々にとって

も、沈黙がコミュニケーションの一部として重要な役割を果たすことが多く [12]、例えば、西アフリカに住むアカン族の間では、沈黙が社会的地位や年齢に対する敬意を示す [13]。

このような沈黙に対する認識の差異は、同じ国の中でも見ることができる。ウェールズの3つの学校の生徒を対象に行った調査では、田舎や都市部の生徒は、校外の生徒に比べて沈黙をより高く評価していることが明らかになり、加えて、生徒と教師の間でも沈黙に対するパワーの不均衡があることが指摘されている [14]。さらに、沈黙の解釈は文脈にも依存し [15–17]、その多義性を容認するための枠組みを提案する試みもある [17]。したがって、このような沈黙の多義性や文化による差異は、特に異文化間のコミュニケーションにおいて深刻なコミュニケーションの誤解を招く可能性がある [3–5]。

沈黙と生産性の関係に着目した研究も数多く存在する。職場での会話における沈黙の重要性に着目した研究では、コミュニケーションのもたらす生産性を議論するためのフレームワークが提案されており、沈黙が生産的な会話にどのように影響するのかが考察されている [18]。また、静寂は人の時間感覚や自己認識に影響を与え、高いリラクゼーションをもたらす場合もある [19]。特に、組織内における沈黙が、従業員のストレスを軽減し、全体的な生産性を向上させ、意思決定プロセスを改善することを示した研究もある [20]。加えて、創造性と沈黙の関係性を調べた研究も存在し、創造的な「内向性」を育むためには、空想にふけるのに役立つような沈黙が必要だとの指摘や [21]、沈黙は新たなアイデアや創造性の発生を促す場となるだけでなく、他者との関係を深化させるための手段であるとの主張もある [22]。

その他、医療現場における沈黙の重要性も認識されている [23, 24]。例えば、思いやりのある沈黙が、患者と医療者の間の理解を促進し、温かい関係を構築するのに役立つことが指摘されている [25]。カウンセリングにおいても、クライアントの沈黙は個人の内面に向き合わせる重要なプロセスであるとの考え方がある [26]。音声記録を用いて、医者と患者の間の対話を分析した研究では、「接続的沈黙」という概念を導入することで、沈黙の違いがコミュニケーションの質にどう影響するかが調べられている [27]。

このように、沈黙は文化的に興味深い性質を持つだけでなく、組織の生産性や医療の質を左右する実用的な側面もある。このような重要性にもかかわらず、その振る舞いをモデルなどを使って解析した例は限定的である。

人の会話を数理モデルとして捉え、その沈黙について調べた初期の研究としては、H. B. Kekre らの研究が挙げられる [28]。彼らは、発話期間と無音期間の二状態を取るマルコフモデルを考え、その分布について考察している [28]。これは、人間の会話における沈黙そのものへの興味というよりはむしろ、当時の限られた通信容量を効率よく利用したいという工学的なニーズによるものだった [28]。2000年代に入ってから、例えば O. I. Sheluhin らによって、電話音声対話のマルコフモデルを利用して通信効率を向上させる提案が行われている [29]。その他、M. Galley らは会議などの参加者の多い会話のセグメンテーション問題に取り組み、語彙的特徴と会話の特徴を統合した手法を提案している [30]。注目すべきは、この会話の特徴付けでは、従来考慮されていなかった沈黙が重要な役割を果たしている点である [30]。いずれも、沈黙の持つ重要な側面を明らかにした研究だが、対象は沈黙そのものではなく、特に集団における沈黙に関する定性的・定量的な振る舞いに関しては明らかにされていない。

一方で、会話を解析するアプローチは、Conversation Analysis(CA) としてまとめられることもあり、非常に多くの研究例がある [31]。A. P. Thomas らの提案した Conversational Exchange Analysis (CEA) は、会話をアクティビティ、タイプ、フォーカスという三つの異なる概念レベルでコーディングし、スピーチのセグ

メンテーションと分類を行うためのルールを提供している [32]。M. Ishizaki と T. Kato は、三者間対話の特性を定量的に調べ、二者間の対話が支配的であることを明らかにした [33]。A. Stolcke らは、n-gram と隠れマルコフモデルを組み合わせ、動的な対話構造を解析する枠組みを提案している [34]。P. M. Aoki らは、機械学習を用いて参加者のターンテイキング行動を分析し、会話のフロアを動的に特定することで、参加者ごとにカスタマイズされた音声配送を実現した [35]。会話を通じて人間の相互作用をモデリングするインフルエンスモデルなども考案されている [36, 37]。K. Otsuka らは、マルコフスイッチングモデルと呼ばれる動的ベイジアンネットワークを使って、会話の構造を推定した [38]。特筆すべきは、視線や頭の向きなどを取り込むことで、非言語的な振る舞いと会話の関係を扱うための確率的フレームワークを提示した点にある [38]。M. Fujimoto と I. Daido は、多人数での会話を主成分分析で解析し、個別の発話行動が、会話の深度や流れにどのように影響するかを調べた [39]。T. Bergstrom と K. Karahalios の提案した Conversation Clock という手法は、会話をしている人間の音声貢献度を基に会話パターンを視覚化し、会話の状況や流れを振り返ることを可能にする [40]。M. Mastrangeli らは、エージェントベースの会話のモデルを提案し、多人数の会話におけるダイナミクスと幸福度の関係を調べている [41]。M. Cristani らは、ガウス混合モデルとマルコ影響モデルを統合し、対話シナリオを分類する枠組みを構築した [42]。D. Angus らは、概念的再現プロットを定義し、用語ベースのアプローチを超えて、任意の二つの発話間の概念的類似性を評価している [43]。彼らは、この方法を医者と患者の会話に適用し、コミュニケーションのパターンと信頼関係の関連を調べている [44]。R. Fusaroli と K. Tylen は、再帰定量化解析と呼ばれる手法を用いて対話を分析し、シナジーが集団のパフォーマンスを予測するための強力な指標であることを示した [45]。L. A. Clarfeld らは、COnversational DYnamics Model (CODYM) という会話の動態や情報の流れを一元的に評価する新たな手法を提案している [46]。これは、スピーカーの発話の長さを基本的な単位としたマルコフモデルであり、これを使って、彼らは医師と重病患者の間に流れる情報の規範的なパターンを明らかにした [46]。また、会話そのもののモデル化ではないが、セラピストとクライアントの感情の変化を 2 次元の常微分方程式で表現した研究があり、パラメータの違いが感情のアトラクターにどのように影響するかが調べられている [47]。その他、互いに相互作用する個人で形成される集団について、その人数が変化することによる社会的構造への影響を調べた研究がある [48]。

このような数多くの会話研究の中で、沈黙が持つ特徴は度々取り上げられてはいるものの、やはりそれ自体を対象として深く解析した研究はない。

そこで本論文では、集団における沈黙の数理モデルを提案し、その振る舞いを解析する。具体的には、会話を構成する人々を、それぞれ沈黙 (0)、発話 (1) の離散的な状態を取るエージェントとし、互いの相互作用を導入したマルコフ連鎖モデルによって会話のダイナミクスを記述する。そして、代表的ないくつかの相互作用状態を詳しく調べることで、沈黙の持つ性質を明らかにすることを試みる。この解析によって、周囲に対する意識を小さくしていくと、あるところで沈黙の発生確率が急激に 0 となる、一種の相転移が観測された。このことから、集団における沈黙、すなわち天使が通るためには、各々がわずかにでも周囲の会話を意識する必要があることが明らかになった。また、本モデルと現実世界との繋がりを議論するため、まだまだ不十分ではあるが、実験とモデルの比較も行った。

本研究は、人々のコミュニケーションにおいて重要な役割を果たす沈黙に着目し、それを解析するための新たな枠組みを提案するものである。これは会話そのもののより深い理解をもたらすだけでなく、異文化交流、組織における生産性、あるいは医療現場などにも寄与できる可能性がある。

2 研究の方法

ここでは、集団における沈黙を解析するための新たな会話モデルを導入する。このモデルでは、沈黙 (0) と発話 (1) の離散的な状態をとるエージェントが複数存在する場を想定する。各エージェントの次の状態は、現在の状態と、エージェント間の相互作用によって決定され、この時間発展の様子を会話と解釈する。これは、H. B. Kekre らのモデルに相互作用を取り入れた自然な拡張だとみなすこともできる。

2.1 1 人の場合

ここで提案するモデルでは、各エージェントは現在の状態を保持する傾向があると仮定する。すなわち、ある時点で沈黙しているエージェントは次の瞬間も沈黙を続ける可能性が高く、逆に発話しているエージェントは次の瞬間も発話を続ける可能性が高い。

ある時刻 t におけるエージェントの状態を $s(t) \in \{0, 1\}$ とし、ここで $s(t) = 0$ は沈黙、 $s(t) = 1$ は発話を表すとする。このとき、 $0 < \alpha, \beta \ll 1$ であるような α, β をそれぞれ沈黙から発話への遷移確率および、発話から沈黙への遷移確率とすれば、先の仮定は次のような状態遷移で記述される。

$$s(t+1) = \begin{cases} 1, & \text{with probability } (1 - \alpha - \beta) s(t) + \alpha \\ 0, & \text{with probability } (\alpha + \beta - 1) s(t) + 1 - \alpha \end{cases}$$

一般に、各ステップ毎に一定の確率 p でその状態が他に遷移するモデルを考えると、その状態が連続する時間の期待値は次の式で与えられる。

$$T = \sum_{n=1}^{\infty} n (1-p)^{n-1} p = \frac{1}{p} \quad (1)$$

これより、先のモデルにおいて、エージェントが沈黙している時間の長さの期待値と、発話している時間の長さの期待値は、それぞれ $1/\alpha, 1/\beta$ で与えられる。定常状態において、沈黙している確率 $f(s=0)$ と、発話している確率 $f(s=1)$ は、それぞれの時間の長さの期待値に比例すると考えてよいので、

$$f(s) = \frac{\alpha}{\alpha + \beta} s + \frac{\beta}{\alpha + \beta} (1 - s)$$

となる。

2.2 $1 \leq N$ 人の場合

前節のモデルを N 人の場合に拡張する。

ここでは、追加の仮定として、エージェントの次の状態は周囲の状況にも影響を受けると考える。すなわち、周りが賑やかであれば、エージェントは発話しやすく、逆に静かであれば沈黙しやすくなる。

これを表現するために、次の量 S を定義する。

$$S_i(t) = \sum_{j=1}^N w_{ij} s_j(t)$$

ここで、 $s_j(t)$ は、ある時刻 t における j 番目のエージェントの状態であり、 w_{ij} は j 番目のエージェントから i 番目のエージェントへの影響の大きさを表す相互作用行列の要素である。また、 w_{ij} は $0 \leq w_{ij}$, $\sum_{j=1}^N w_{ij} = 1$ を満たすものとする。

これを使うと、 i 番目のエージェントの状態遷移は次のように記述できる。

$$s_i(t+1) = \begin{cases} 1, & \text{with probability } (1 - \alpha - \beta) S_i(t) + \alpha \\ 0, & \text{with probability } (\alpha + \beta - 1) S_i(t) + 1 - \alpha \end{cases} \quad (2)$$

もし、 $w_{ij} = \delta_{ij}$ であれば、これは周囲の状態を一切無視し、次の状態が今の自分の状態にしか依存しない場合となり、前節のモデルと完全に一致する。

この w_{ij} は、このモデルにおける会話を特徴付ける。例えば、同じ会話グループに属している人間同士でも、互いに同じ影響を与え合うとは限らない。同じ空間であっても、例えば、生徒と教師の間には沈黙に対するパワーの不均衡があることが指摘されている [14]。また、家族の会話を調べた研究では、会話の頻度が参加者間の社会的距離や権力に影響を受けることが示されており、さらに特定の状況では、これによって応答が意図的に制限される [49]。このような不均衡の最も簡単な導入は、 w を非対称行列にすることによって実現される。

2.3 確率分布

N 人の場合について、ある時刻 t に話している人数の確率分布を考える。

2.3.1 孤立集団

ここでは、 $w_{ij} = \delta_{ij}$ の場合、すなわち、各エージェントが周囲の状況を一切無視するような場合を考える。

このとき、十分時間が経った後で n_1 人が話している確率 $f_{\text{iso}}(n_1)$ は、二項分布から

$$f_{\text{iso}}(n_1) = \binom{N}{n_1} \left(\frac{\alpha}{\alpha + \beta} \right)^{n_1} \left(\frac{\beta}{\alpha + \beta} \right)^{N-n_1}$$

となる。

2.3.2 均一集団

ここでは、 $w_{ij} = 1/N$ の場合を考える。これは、集団を構成する各エージェントが、全員 (自分を含めた) から等しく影響を受けているような状況であり、次の状態への遷移確率は、エージェントに依らず同じ形式で書ける。

このとき、ある時刻 t に n_0 人が沈黙、 n_1 人が話していて、次の瞬間 (時刻 $t + \Delta t$) に、新たに d_{01} 人が話し始め、 d_{10} 人が沈黙する場合の確率 P は、

$$P(d_{01}, d_{10}; N, n_0, n_1) = \binom{n_0}{d_{01}} \binom{n_1}{d_{10}} p_{00}^{n_0-d_{01}} p_{01}^{d_{01}} p_{10}^{d_{10}} p_{11}^{n_1-d_{10}}$$

と書ける。ここで、 p_{uv} は、状態 u から状態 v への遷移確率である。それぞれの確率は、(2) 式より

$$\begin{aligned} p_{00} &= p_{10} = \frac{n_0}{N} (1 - \alpha) + \frac{n_1}{N} \beta \\ p_{01} &= p_{11} = \frac{n_0}{N} \alpha + \frac{n_1}{N} (1 - \beta) \end{aligned}$$

となる。

以上より、話している人数が、時刻 t から $t + \Delta t$ で、 n_1 人から n'_1 人になる確率 $g(n_1, n'_1)$ は、

$$g(n_1, n'_1) = \sum_{(d_{01}, d_{10}) \in D} \binom{n_0}{d_{01}} \binom{n_1}{d_{10}} \left(\frac{n_0}{N} \alpha + \frac{n_1}{N} (1 - \beta) \right)^{n'_1} \left(\frac{n_0}{N} (1 - \alpha) + \frac{n_1}{N} \beta \right)^{N-n'_1}$$

(where $D = \{(d_{01}, d_{10}) \mid d_{01} - d_{10} = n'_1 - n_1\}$)

と計算できる。ここで、 D は、 $d_{01} - d_{10} = n'_1 - n_1$ を満たす (d_{01}, d_{10}) のペアの集合である。

この総和を計算すると

$$g(n_1, n'_1) = \binom{N}{n'_1} \left(\frac{n_0}{N} \alpha + \frac{n_1}{N} (1 - \beta) \right)^{n'_1} \left(\frac{n_0}{N} (1 - \alpha) + \frac{n_1}{N} \beta \right)^{N-n'_1}$$

を得る。

途中で

$$\sum_{(d_{01}, d_{10}) \in D} \binom{n_0}{d_{01}} \binom{n_1}{d_{10}} = \binom{N}{n'_1}$$

を使った。

これより、十分時間が経った後で n_1 人が話している確率 $f_{\text{uni}}(n_1)$ が満たすべき方程式として

$$f_{\text{uni}}(n'_1) = \sum_{n_1=0}^N g(n_1, n'_1) f_{\text{uni}}(n_1)$$

を得る。

一般にこれを解くことは難しい。しかし、これを積分方程式で書き直し、 $0 < \alpha, \beta \ll 1$ を使うと、この近似的な解として

$$f_{\text{uni}}(n_1) = \frac{1}{Z(2\alpha N, 2\beta N)} \left\{ \left(1 - \frac{n_1}{N}\right) \alpha + \frac{n_1}{N} (1 - \beta) \right\}^{2\alpha N-1} \left\{ \left(1 - \frac{n_1}{N}\right) (1 - \alpha) + \frac{n_1}{N} \beta \right\}^{2\beta N-1}$$

を得る。ここで、

$$Z(2\alpha N, 2\beta N) = \sum_{n_1=0}^N \left\{ \left(1 - \frac{n_1}{N}\right) \alpha + \frac{n_1}{N} (1 - \beta) \right\}^{2\alpha N - 1} \left\{ \left(1 - \frac{n_1}{N}\right) (1 - \alpha) + \frac{n_1}{N} \beta \right\}^{2\beta N - 1}$$

である。

この詳しい導出は Appendix に記した。

2.3.3 喫茶店 θ モデル

ここでは、エージェントの次の状態を決定する要因を、現時刻における「そのエージェントの状態」と、「環境の状態」の2つに明示的に切り分けることを考える。

まず、 $0 \leq \theta \leq 1$ を満たすパラメータ θ を使って、 w_{ij} を次のように定義する。

$$w_{ij} = \begin{cases} \theta & (i = j) \\ \frac{1-\theta}{N-1} & (i \neq j) \end{cases} \quad (3)$$

この θ は、次のエージェントの状態が、今のエージェント自身の状態にどれだけ影響を受けるかを示すものであり、 $\theta = 0$ の場合は、次の状態が周辺の環境によってのみ決定され、 $\theta = 1$ の場合には、次の状態は今のエージェント自身の状態にのみ依存する。

このような w_{ij} において、十分時間が経ったときに、 n_1 人が話している確率 $f_\theta(n_1)$ を考える。また、このとき N は十分大きいとする。

$\theta = 0$ の場合、この w_{ij} は、 $w_{ij} \sim 1/N$ とみなすことができる。したがって、これは2.3.2の結果から

$$f_{\theta=0}(n_1) = \frac{1}{Z(2\alpha N, 2\beta N)} \left\{ \left(1 - \frac{n_1}{N}\right) \alpha + \frac{n_1}{N} (1 - \beta) \right\}^{2\alpha N - 1} \left\{ \left(1 - \frac{n_1}{N}\right) (1 - \alpha) + \frac{n_1}{N} \beta \right\}^{2\beta N - 1}$$

となる。

次に $\theta = 1$ の場合には、 $w_{ij} = \delta_{ij}$ となるため、これは2.1の結果と完全に一致し、

$$f_{\theta=1}(n_1) = \binom{N}{n_1} \left(\frac{\alpha}{\alpha + \beta} \right)^{n_1} \left(\frac{\beta}{\alpha + \beta} \right)^{N - n_1}$$

となる。

ここで、 N が十分大きく、また、 $1 \ll \alpha N / (\alpha + \beta)$ であるので、この二項分布は正規分布に近似できる。

$$\begin{aligned} f_{\theta=1}(n_1) &= \binom{N}{n_1} \left(\frac{\alpha}{\alpha + \beta} \right)^{n_1} \left(\frac{\beta}{\alpha + \beta} \right)^{N - n_1} \\ &= \binom{N}{n_1} \left(\frac{\alpha}{\alpha + \beta} \right)^{n_1} \left(1 - \frac{\alpha}{\alpha + \beta} \right)^{N - n_1} \\ &\sim \frac{1}{\sqrt{2\pi \frac{\alpha\beta N}{(\alpha + \beta)^2}}} \exp \left(-\frac{\left(n_1 - \frac{\alpha N}{\alpha + \beta} \right)^2}{2 \frac{\alpha\beta N}{(\alpha + \beta)^2}} \right) \end{aligned}$$

さらに、 $n_1 \rightarrow n_1/N$ とスケーリングさせれば、これはベータ分布に近似できる。

$$f_{\theta=1}(n_1) \sim \frac{1}{\sqrt{2\pi \frac{\alpha\beta}{(\alpha+\beta)^2} \frac{1}{N}}} \exp\left(-\frac{\left(\frac{n_1}{N} - \frac{\alpha}{\alpha+\beta}\right)^2}{2 \frac{\alpha\beta}{(\alpha+\beta)^2} \frac{1}{N}}\right) \\ \sim B\left(\frac{\alpha N}{\alpha+\beta}, \frac{\beta N}{\alpha+\beta}\right)$$

また、 $0 < \alpha, \beta \ll 1$ であるので

$$\frac{n_1}{N} \sim \left(1 - \frac{n_1}{N}\right) \alpha + \frac{n_1}{N} (1 - \beta) \\ 1 - \frac{n_1}{N} \sim \left(1 - \frac{n_1}{N}\right) (1 - \alpha) + \frac{n_1}{N} \beta$$

とすれば、結局

$$f_{\theta=1}(n_1) \sim \frac{1}{Z\left(\frac{\alpha N}{\alpha+\beta}, \frac{\beta N}{\alpha+\beta}\right)} \left\{\left(1 - \frac{n_1}{N}\right) \alpha + \frac{n_1}{N} (1 - \beta)\right\}^{\frac{\alpha N}{\alpha+\beta}-1} \left\{\left(1 - \frac{n_1}{N}\right) (1 - \alpha) + \frac{n_1}{N} \beta\right\}^{\frac{\beta N}{\alpha+\beta}-1}$$

となる。

以上の結果から、一般的な θ について、 f_θ の形を次のように仮定する。

$$f_\theta(n_1) = \frac{1}{Z(h_\alpha(\theta), h_\beta(\theta))} \left\{\left(1 - \frac{n_1}{N}\right) \alpha + \frac{n_1}{N} (1 - \beta)\right\}^{h_\alpha(\theta)-1} \left\{\left(1 - \frac{n_1}{N}\right) (1 - \alpha) + \frac{n_1}{N} \beta\right\}^{h_\beta(\theta)-1} \quad (4)$$

ここで、 $h_\alpha(\theta), h_\beta(\theta)$ は、それぞれ次の境界条件を満たす関数である。

| θ | 0 | \rightarrow | 1 |
|--------------------|-------------|---------------|----------------------------------|
| $h_\alpha(\theta)$ | $2\alpha N$ | | $\frac{2\alpha N}{\alpha+\beta}$ |
| $h_\beta(\theta)$ | $2\beta N$ | | $\frac{2\beta N}{\alpha+\beta}$ |

表1: Boundary conditions for $h_\alpha(\theta)$ and $h_\beta(\theta)$.

ここで導出した (4) 式は、 N が十分に大きいと仮定している。したがって、これは 1 対 1 の会話で成り立つような分布ではなく、喫茶店やレストラン、図書館や教室などの不特定多数の人間が同じ空間に存在している場合にうまく適合するモデルだと考えられる。そこで、本論文ではこれを喫茶店 θ モデルと呼ぶことにする。

3 結果と考察

以降では、喫茶店 θ モデルを使って議論を進める。

3.1 補間関数 h

ここでは、補間関数 $h_\alpha(\theta), h_\beta(\theta)$ を、シミュレーションの結果に合うように決定する。

まずは、境界条件を満たすように、補間関数の形を次のように仮定する。

$$\begin{aligned} h_\alpha(\theta) &= N\alpha \left(2(1 - h(\theta)) + \frac{h(\theta)}{\alpha + \beta} \right) \\ h_\beta(\theta) &= N\beta \left(2(1 - h(\theta)) + \frac{h(\theta)}{\alpha + \beta} \right) \end{aligned} \quad (5)$$

ここで、 $h(\theta)$ は $\theta = 0$ で 0、 $\theta = 1$ で 1 を満たすような関数であり、簡単のため頭には書いていないが、実際には各パラメータ (N, α, β) にも依存する。この依存性を確認するために、様々なパラメータの組合せでシミュレーションを行い、そのときの尤もらしい h の値を調べることにする。具体的には、各パラメータに対して、シミュレーションによって得られた確率分布と、(4) 式との Kullback-Leibler(KL) ダイバージェンスを評価し、それを最小にするような h_{\min} を、尤もらしい h とする。

今回は、Table2に示すパラメータの組合せでシミュレーションを行った。

表2: シミュレーションで使したパラメータ設定

| パラメータ | 値の範囲 |
|----------|--|
| N | 512, 1024, 2048, 4096 |
| α | $\frac{1}{100}, \frac{1}{200}, \frac{1}{400}, \frac{1}{800}, \frac{1}{1600}, \frac{1}{3200}, \frac{1}{6400}$ |
| β | $\frac{1}{100}, \frac{1}{200}, \frac{1}{400}, \frac{1}{800}, \frac{1}{1600}, \frac{1}{3200}, \frac{1}{6400}$ |
| θ | 0.0, 0.25, 0.50, 0.75, 0.80, 0.85, 0.90, 0.95, 1.0 |

各パラメータを変えたときの、 h_{\min} の変化を Fig. 1aに示す。 h_{\min} と θ の間に正の相関がある以外には、ほとんど相関が見られないことが分かる。

ここで、シミュレーションで得られた各パラメータと、それに対する h_{\min} の組に対して PySR を使ってシンボリック回帰を行った [50]。提案されたいくつかの回帰モデルの中から、境界条件を適切に満たすシンプルなものを選ぶと次のようになった。

$$h(\theta) \sim \frac{\theta(\alpha + \beta)}{1 - \theta + \alpha + \beta} \quad (6)$$

シミュレーションの結果と、この回帰式の結果を比較したプロットを Fig. 1bに示す。非常に良く一致していることが確認できる。

この結果を踏まえ、以降では、補間関数 $h_\alpha(\theta), h_\beta(\theta)$ には、(5) 式の仮定と、(6) 式の近似を採用する。

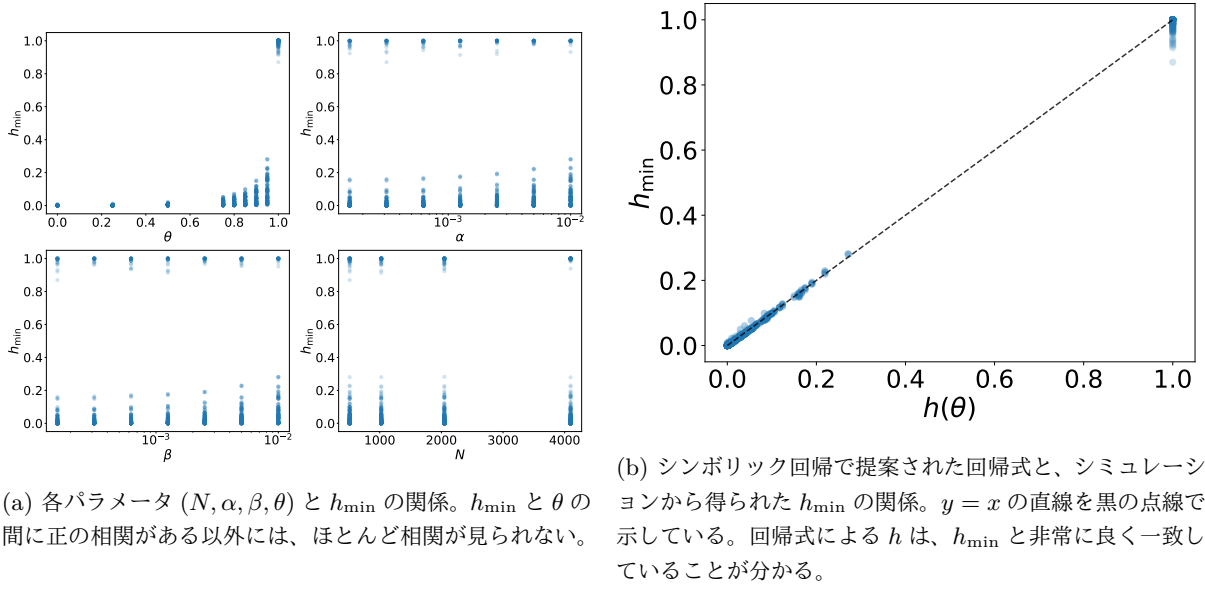


図1: 様々なパラメータ、回帰式と h_{\min} の関係。

3.2 近似の評価

ここでは、喫茶店 θ モデルとシミュレーションの結果を比較することで、導出の際に行った近似が妥当であるかを検証する。

まずは、 N が大きい場合に両者が一致していることを確認するために $N = 512$ の結果を Fig. 2 に示す。どのパラメータについても、喫茶店 θ モデルとシミュレーションの結果は非常に良く一致していることが確認できる。

次に、この一致が成り立つ N を見極めるために、 $N = 4, 8, 64$ の場合の結果を Fig. 3 に示す。喫茶店 θ モデルの導出過程で、 N が大きいという仮定をおいたにもかかわらず、 $N = 4$ という N が非常に小さい場合でも、近似式とシミュレーションの結果はよく一致していることが分かる。

3.3 ランダムな重み w と喫茶店 θ モデル

ここでは、重み w をランダムにした場合に、喫茶店 θ モデルからどれだけ乖離するのかを調べる。

重み w の非対角成分は、周辺からどれだけ影響を受けるかを示すものである。したがって、 w_{ij} は、 i 番目と j 番目のエージェントの距離が離れば離れるほど小さくなるように決めるのが尤もらしい。ここでは、まず N 人のエージェントを 2 次元平面上にランダムに配置し、 w'_{ij} を距離の二乗に反比例する値として決める(対角成分は 0 とする)。そして、この w' を使って、 w を次のように決定する。

$$w_{ij} = \begin{cases} \theta & (i = j) \\ (1 - \theta) \frac{w'_{ij}}{\sum_k w'_{ik}} & (i \neq j) \end{cases}$$

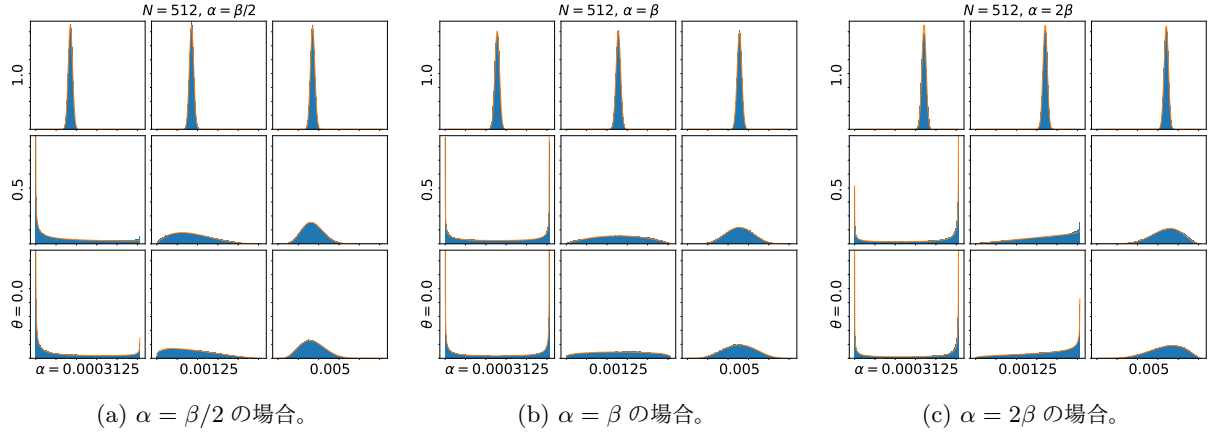


図2: 喫茶店 θ モデル (オレンジの実線) とシミュレーション (青のヒストグラム) の比較。 α, β の関係を 3 パターン用意し、それぞれの場合で、 $\theta = 0.0, 0.5, 1.0$, $\alpha = 1/3200, 1/800, 1/200$ でパラメータを変化させたときの分布を示している。(a) は $\alpha = \beta/2$, (b) は $\alpha = \beta$, (c) は $\alpha = 2\beta$ の結果。どのパラメータも、両者の結果は非常によく一致していることが分かる。

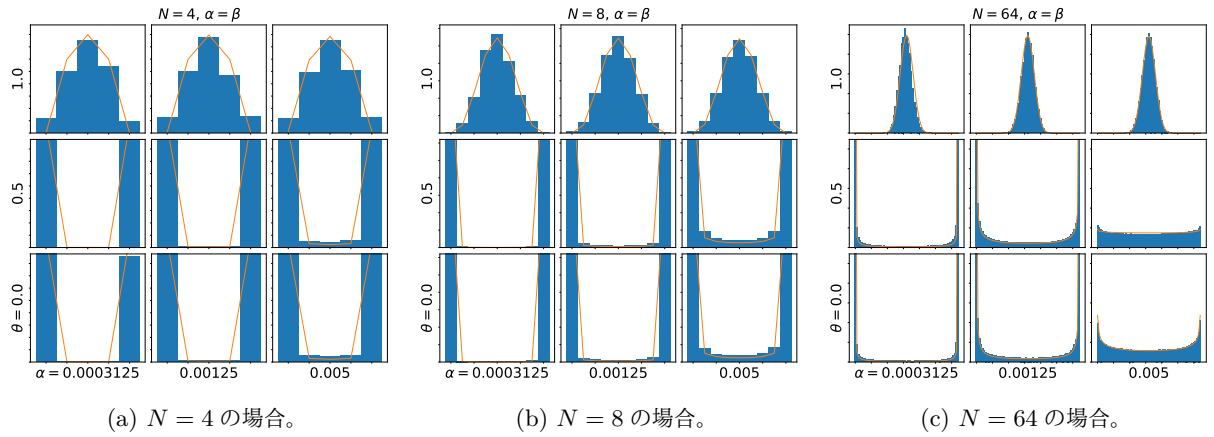


図3: 喫茶店 θ モデル (オレンジの実線) とシミュレーション (青のヒストグラム) の比較。 $\alpha = \beta$ の場合のみを考え、 $\theta = 0.0, 0.5, 1.0$, $\alpha = 1/3200, 1/800, 1/200$ でパラメータを変化させたときの分布を示している。(a) は $N = 4$, (b) は $N = 8$, (c) は $N = 64$ の結果。どのパラメータも、両者の結果は非常によく一致していることが分かる。

N 人の座標は、 $[0, 1)$ の一様乱数によって決定し、各パラメータについて 10 パターン用意した。

まずは N が大きい場合の例として、 $N = 512$ の結果を Fig. 4 に示す。重みを乱数によって決定しているにもかかわらず、喫茶店 θ モデルとシミュレーションの結果は驚くほど一致していることが分かる。

次に、 N が小さい場合の例として、 $N = 4, 8, 64$ の結果を Fig. 5 に示す。 N が小さい場合でも、喫茶店 θ モデルで大まかな傾向を捉えられていることが分かる。一方で、例えば Fig. 5a に多く見られるような、中央と両端で度数が大きくなるような分布や、Fig. 5b の $\alpha = 1/3200, w = 0.95$ で見られるような多峰性の分布の形

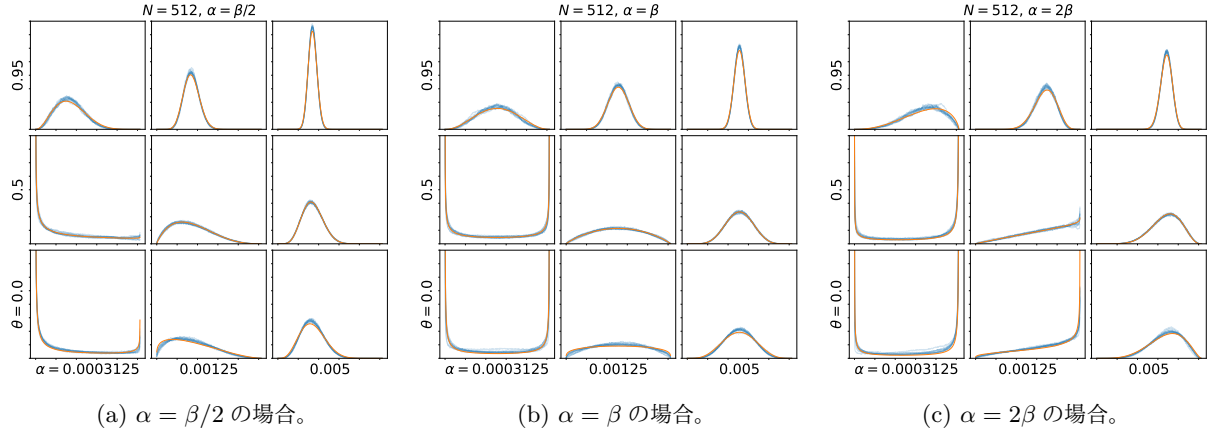


図4: 喫茶店 θ モデル (橙色) とシミュレーション (青) の比較。シミュレーションの方は、乱数を変えた 10 パターンの結果を重ねてプロットしている。 α, β の関係を 3 パターン用意し、それぞれの場合で、 $\theta = 0.0, 0.5, 0.95$, $\alpha = 1/3200, 1/800, 1/200$ でパラメータを変化させたときの分布を示している。(a) は $\alpha = \beta/2$, (b) は $\alpha = \beta$, (c) は $\alpha = 2\beta$ の結果。どのパラメータも、両者の結果は非常によく一致していることが分かる。

状を捉えられていないことが分かる。

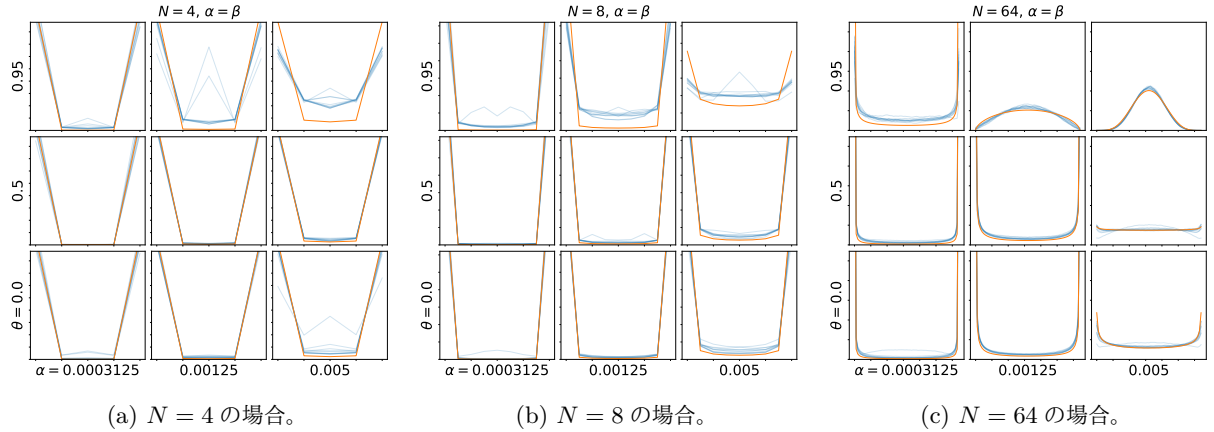


図5: 喫茶店 θ モデル (橙色) とシミュレーション (青) の比較。シミュレーションの方は、乱数を変えた 10 パターンの結果を重ねてプロットしている。 $\alpha = \beta$ の場合のみを考え、 $\theta = 0.0, 0.5, 0.95$, $\alpha = 1/3200, 1/800, 1/200$ でパラメータを変化させたときの分布を示している。(a) は $N = 4$, (b) は $N = 8$, (c) は $N = 64$ の結果。どのパラメータも、両者の大まかな傾向はよく一致しているが、多峰性の分布の形状を捉えられていない例があることが分かる。

3.4 沈黙の性質

ここでは、喫茶店 θ モデルにおいて沈黙が発生する条件や、その持続時間などについて考察する。

まず、 α, β の意味を改めて確認すると、これらはそれぞれ、沈黙から発話への遷移確率と、発話から沈黙へ

の遷移確率として導入されたものである。したがって、 α を大きくすると、全体の分布は発話寄りとなり、 β を大きくすると、全体の傾向は沈黙寄りとなる。

今興味があるのは、そのような自明な傾向ではなく、系全体としては一定の発話量がある状態での沈黙の挙動である。そこで、ここでは話している人数の期待値 $\langle n_1 \rangle$ が $N/2$ であるような場合について考える。

(4) 式から、期待値 $\langle n_1 \rangle$ は

$$\langle n_1 \rangle = N \frac{h_\alpha}{h_\alpha + h_\beta} = N \frac{\alpha}{\alpha + \beta}$$

と、 θ に依らない表式となる。これより、 $\langle n_1 \rangle = N/2$ となるのは、 $\alpha = \beta$ の場合であることが分かる。

この条件で、沈黙が発生する確率 $f_\theta(n_1 = 0) = P_0$ を考える。いくつかの N について、その他のパラメータ (α, β, θ) を変えたときの P_0 をヒートマップとして Fig. 6 に示した。

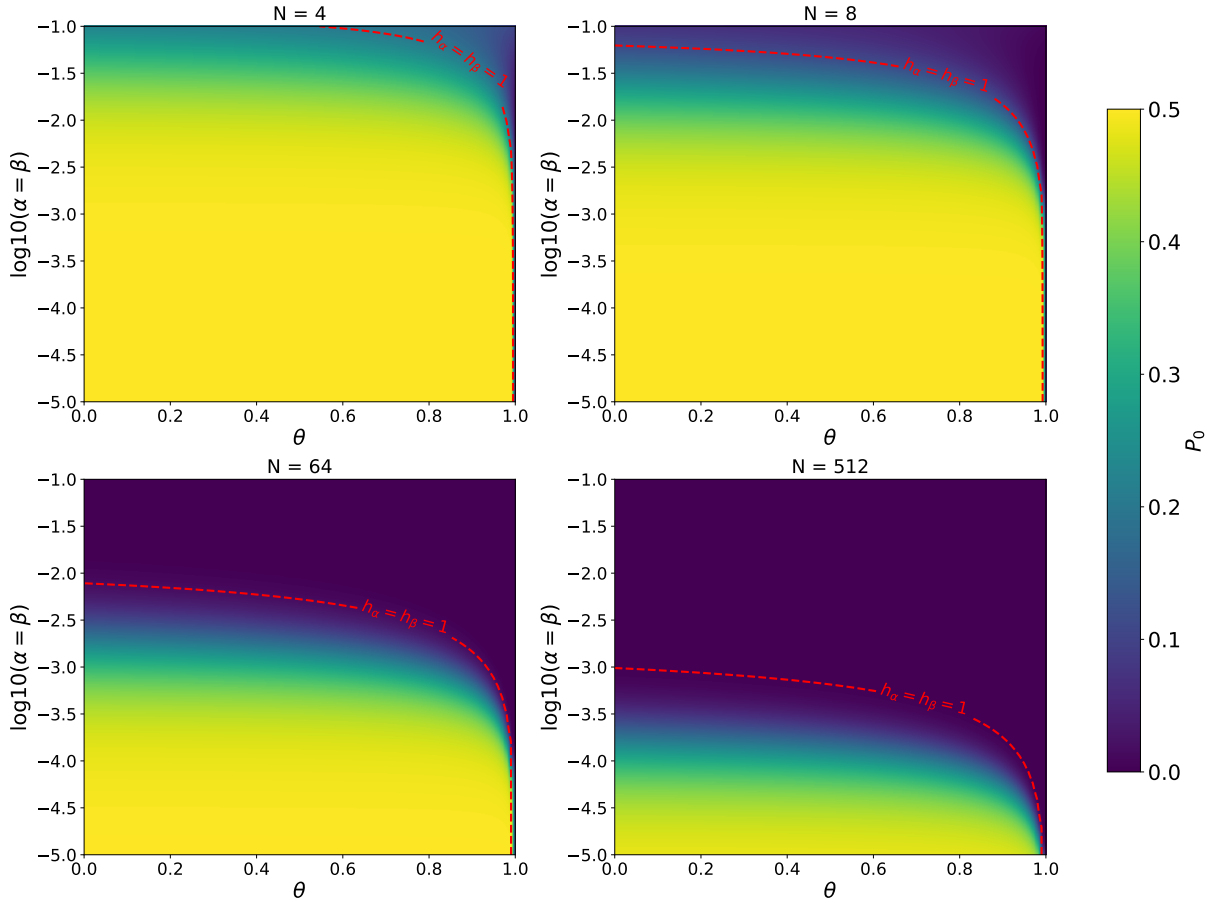


図6: $N = 4, 8, 64, 512$ の場合の、パラメータ (α, β, θ) に対する沈黙の発生確率 P_0 の変化。 $h_\alpha = h_\beta = 1$ の点線は、 f_θ が一様分布になる部分を示し、ここを境に P_0 の振る舞いは劇的に変化する。

発話への遷移確率 α と、沈黙への遷移確率 β が小さい場合には、 P_0 が高くなることが分かる。これはすなわち、現状を維持する傾向が強い場合には沈黙が発生しやすいことを意味する。また、 N が大きくなると、 P_0

がほとんど 0 となる領域が広がっていくことが分かる。これは、人数が増えれば沈黙が発生しづらくなることに相当し、これらは直感にも合致している。

次に θ に着目する。 θ を大きくしていくと、 $h_\alpha = h_\beta = 1$ となるところで P_0 が急激に 0 となることが分かる。この相転移のような現象は、 f_θ がちょうど一様分布になるところで発生しており、ここを境に f_θ の振る舞いは劇的に変化する。 $h_\alpha = h_\beta < 1$ となる領域では、分布が $f_\theta(n_1 = 0)$ と $f_\theta(n_1 = N)$ に均等に偏るため、ほぼ全員が発話している状態と、ほぼ全員が沈黙している状態を行き来するようなダイナミクスに対応する。一方、 $1 < h_\alpha = h_\beta$ の領域では、 f_θ は正規分布のようになり、 P_0 はほとんど 0 である。言い換えれば、 θ は、自分に対する自分の影響度を示すものであるため、これを大きくすることで、各エージェントは周りの環境にほとんど影響されないようになり、このような状況では沈黙は発生しづらい。これより、集団における沈黙は、各個人が周囲を意識して初めて発生するものであることが示唆される。

最後に、発生した沈黙の継続時間 τ を考える。場にいる全員が沈黙している状態から、次の瞬間その沈黙が破られる確率、すなわち少なくとも 1 人が話し出す確率は、余事象から $1 - (1 - \alpha)^N$ と書くことができ、これは θ に依存しない。ここで (1) 式を使えば、 τ の期待値として

$$\langle \tau \rangle = \frac{1}{1 - (1 - \alpha)^N} \sim \frac{1}{\alpha N}$$

を得る。こちらについては、沈黙から発話への遷移確率が低く、また人数が少ない時に沈黙が持続しやすいという直感とも一致した結果を得る。

この $\langle \tau \rangle$ について、上式とシミュレーションにおける τ の平均値を比較したプロットを Fig. 7 に示した。そもその沈黙の発生確率 P_0 が低い場合には、シミュレーションでは $\langle \tau \rangle$ を正確に評価できず 0 となってしまう。そこで、 $P_0 = 10^{-3}$ を境目として、左右でプロットを分割した。左側の P_0 が十分大きい場合の比較では、 $\langle \tau \rangle$ の理論値とシミュレーションから得られた τ の平均値が良く一致していることが分かる。ここで、 $\langle \tau \rangle$ が大きいところでは、そのばらつきも大きくなっていることが分かる。これは τ の分散 $\text{Var}(\tau)$ が

$$\text{Var}(\tau) = \frac{(1 - \alpha)^N}{(1 - (1 - \alpha)^N)^2} \sim \frac{(1 - \alpha)^N}{\alpha^2 N^2}$$

と書けることから理解できる。

3.5 実験との比較

ここでは、喫茶店 θ モデルを CHiME-6 データセットと比較する。これは、複数の話者が自然に会話しているときの音声データを話者ごとに分離するという課題のために作られたデータセットであり、ここには実際の家で行われた 4 人のディナーパーティの様子が録音されている [51]。

喫茶店 θ モデルは N が十分大きいという仮定の下で導出されているため、この 4 人のデータとの比較はあまり適切ではないことを強調しておきたい。しかし、このデータは、我々が入手できた中で最も話者が多く、かつ自然な会話であったため、このモデルを評価するための最初のステップとして比較を試みた。将来的には、より多くの人数が参加した実験データが検証のために使用されるべきである。

このような事情から、喫茶店 θ モデルと CHiME-6 データセットを比較するには工夫が必要である。このデ

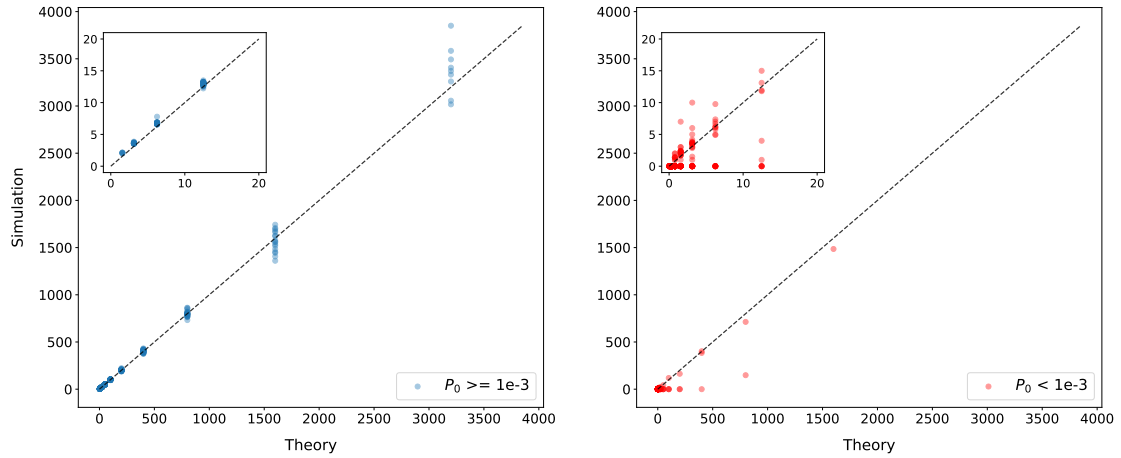


図7: 発生した沈黙の継続時間 τ について、横軸にその理論的な期待値をとり、縦軸にシミュレーションから得られた平均値をプロットした。単位はステップ数。そもそもの沈黙の発生確率 P_0 が低い場合には、シミュレーションで τ の平均値を正確に評価できないため、 $P_0 = 10^{-3}$ を境目として、左右でプロットを分割した。左側の P_0 が十分大きい場合の比較では、理論値とシミュレーション結果が良く一致している。

ータセットを確認してまず気がつくのは、全ての話者が常に同じ会話グループに属しているということである。すなわち、ここでは会話グループの分裂が発生しない。したがって、ある時刻に発話している話者は基本的に1人だけである。一方、我々のモデルが想定しているのは、文字通り喫茶店のような環境であり、複数の会話グループで同時に別々の会話が進行し、それゆえ複数の話者が同時に存在することが自然な状況である。したがって、CHiME-6 では会話の盛り上がり (場の声量の大きさ) は、あくまで個人の声量の大きさによってもたらされるのに対して、喫茶店 θ モデルでは、場の盛り上がりは、そのときに発話している話者の人数によってもたらされる。

そこで、この差を埋めるために、喫茶店 θ モデルを次のように解釈し直す。まず、複数体のエージェントをまとめて1体のエージェントだと考えてみる。また、構成要素のエージェントを子-エージェントと呼び、それらをまとめて1つにしたものを親-エージェントと呼ぶことにする。そして、発話している子-エージェントの数を、親エージェントの声量だとみなす。したがって、各親-エージェントを構成する子-エージェントの数をそれぞれ M とすると、親-エージェントの声量は最小で0から最大で M までの解像度を持つことになる。ここでは、4人の話者に対し、単に $4M$ 人の喫茶店 θ モデルを対応させることで、これを表現することとした。また、今回の検証では $M = 8$ としたが、これは M が8以上であれば、だいたい同じような結果になることを断っておく。

今回の検証では、喫茶店 θ モデルと比較すべき分布を、実際の音声データから次の手順で算出した。まず、フレーム長 512 サンプル、ホップ長 256 サンプルで短時間フーリエ変換 (STFT) を行い、パワースペクトラムから各フレームのエネルギーを算出した。このとき、エネルギーが0のデータは取り除いた。このエネルギーが0の部分は、沈黙と言うより、何らかの理由で全く録音がなされていない区間を指す。その後、デシベルスケールに変換し、外れ値と思われるものを除去し、その分布を比較対象とした。

CHiME-6 には、24 セッション分の録音データが含まれており、また、それぞれ録音場所によってさらに複

数のユニットに別れる。各データの録音時間はおおよそ 2 時間となっており、上記の分布を確認したところ、この 2 時間の間に、分布の傾向が大きく変わっているデータがあることが分かった。同じセッション内で分布の顕著な推移が見られた例を Fig. 8 の上図に示す。このようなダイナミックな分布の変化は、異なる性質の会話に移行した結果だとみなすべきである。今回のモデルで言えば、 α や β 、あるいは w が変化したと考える必要がある。したがって今回の検証では、このような分布の移行が見られないデータを選び比較を行った。具体的には、1 つの音声データを 10 個の時間領域に分割し、各領域におけるデシベルスケールの最小値を取得、この分散が小さいものを選ぶことにした。この分散が最も小さかったデータを Fig. 8 の下図に示す。

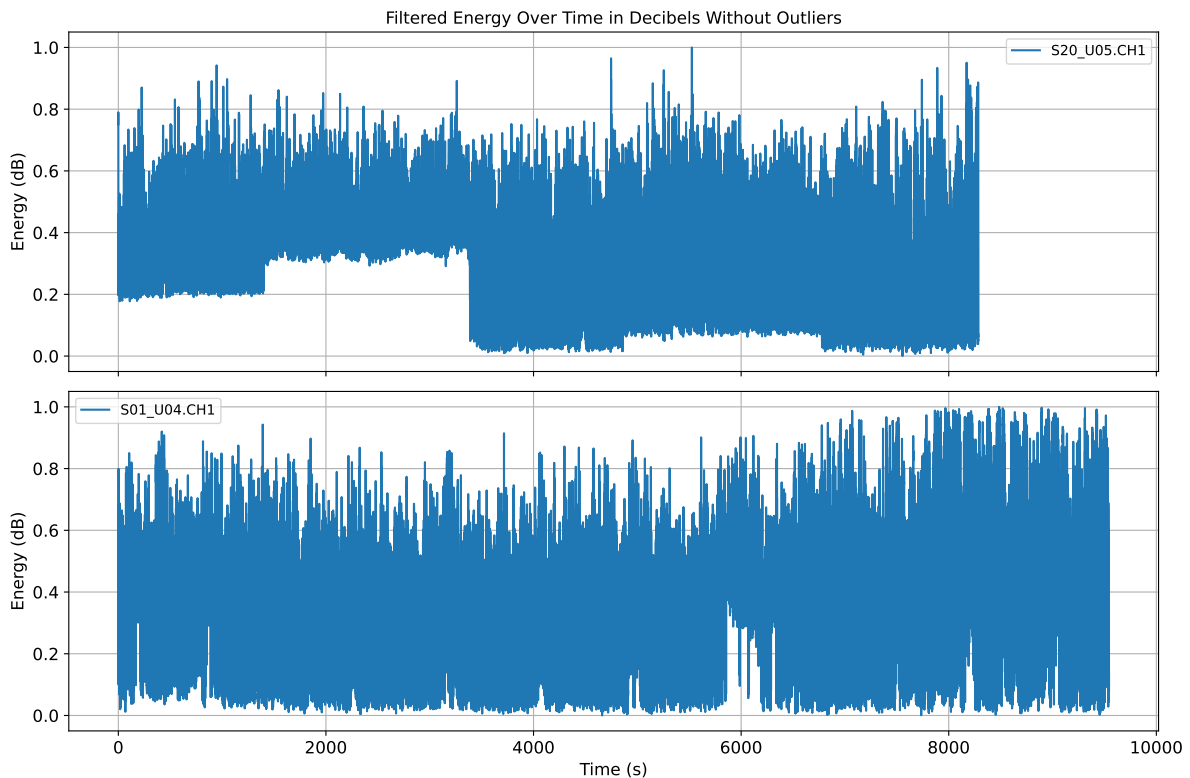


図8: CHiME-6 において、パワースペクトラムから得られた各フレームのエネルギー (デシベルスケール) の分布。分布の傾向が大きく推移する例を上図に、分布の傾向が一定である例を下図に示した。

分散の小さい上位 9 つの音声データと、喫茶店 θ モデルを比較した結果を Fig. 9 に示す。喫茶店 θ モデルのパラメータは、 $h = 0$ とし、 α, β をモーメント法によって決定した。

大まかな分布は捉えられているものの、一部の音声データは多峰性の分布を持っており、単純なベータ分布で表現される喫茶店 θ モデルでは、今回のような 4 人の会話の特徴を表現しきれていないことが分かる。興味深いことに、この多峰性はランダムな重みを与えた場合の結果 (Fig. 5) と整合する。

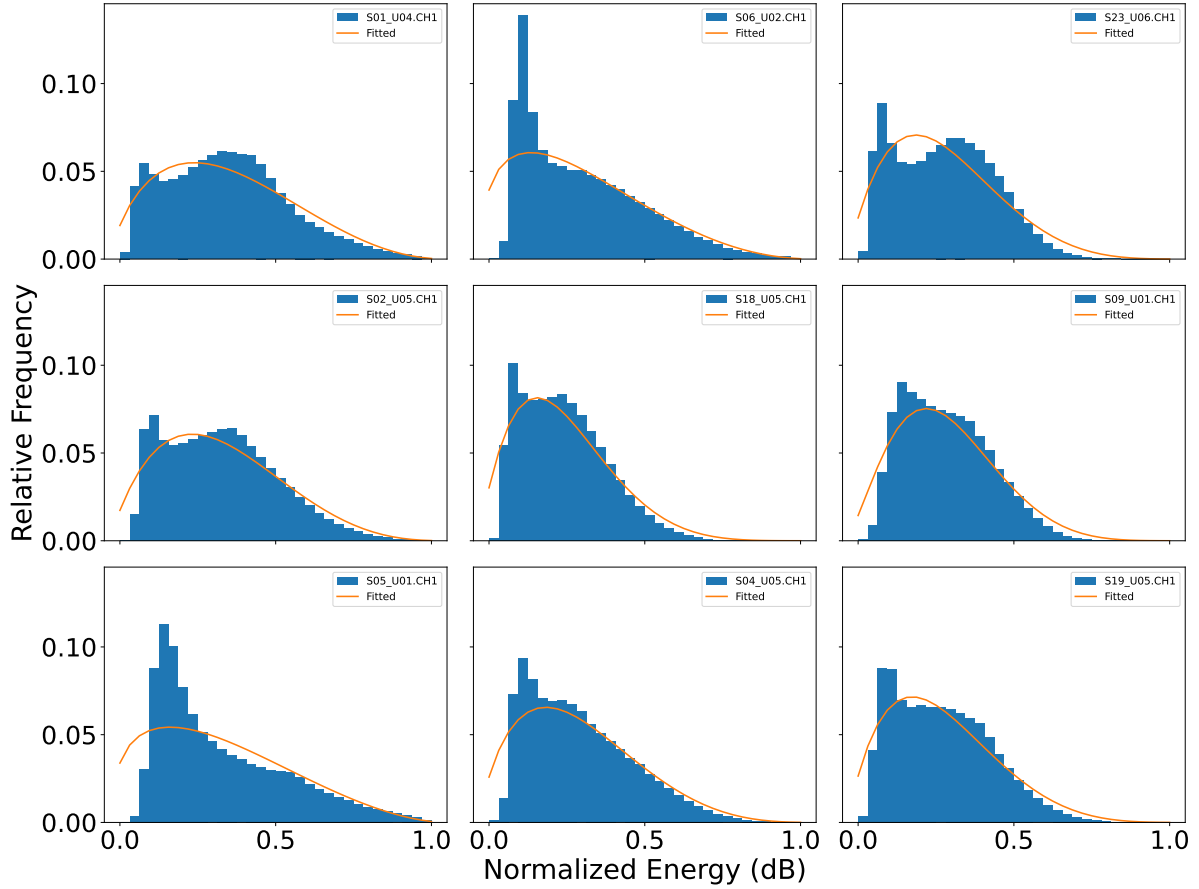


図9: 分散の小さい CHiME-6 のデータ 9 つと、喫茶店 θ モデルの比較。モデルは、大まかな分布は捉えられているものの、一部の音声データが持つ多峰性を表現できていない。

3.6 本提案モデルの課題と限界

ここでは、実際の会話を持つ特徴を文献から調べることで、今回提案したマルコフ連鎖モデルの課題と限界を明らかにする。

まず、同じ発話という行為でも、会話の内容によって聞き手の発話に与える影響が異なることは想像に難しくなく、さらに、「uh huh」などの単語的意味を持たない応答行動が、リスナーの会話への積極性に影響を与えることもある [52]。加えて、参加者の中で生じる小さな会話グループに着目した研究では、グループの分裂が特定の発言によって引き起こされていることなどが明らかになっている [53, 54]。本論文で提案されたモデルは、発話行為を全て 1 という画一的なラベルで置き換えてしまっており、このような効果を取り入れることができない。

また、会話の研究の長い歴史の中で、非言語的なコミュニケーションが会話に影響を与えることは、繰り返し強調されてきた [40, 44, 54–65]。参加者は相互の反応を観察しながら、発話を調整し、会話の交代は、表情やうなずき、ジェスチャーなどによっても引き起こされる [55, 58, 63–65]。また、この際、身体的距離や身体

の向きも重要な要素となる [60]。さらに、視覚的な情報は、沈黙の発生にも大きな影響を与える [66]。しかしながら、本モデルではこれらの要素が完全に無視されている。特に発話の重複については、実際には発生後速やかに解決されるのに対して [67]、我々のモデルではむしろそれが強調される。

このような現実との差異は、例えば、モデルが取り得る状態をより細かく分類し、また、複数の相互作用の存在を許すような拡張を行うことで、ある程度は緩和できるかもしれない。実際、沈黙の種類を分類した非常に多くの研究存在し [13, 25, 68–72]、当然、発話の種類を分類する試みも古くから行われている [32, 56]。特に、相互作用に負の値を許すことで、ある発話がまた別の発話を抑制する状況を表現できるようになるため、重複の解消とターニングの再現が期待される。

一方で、沈黙には、本モデルを単純に拡張しただけでは取り扱うのが難しい側面も存在する。例えば、実際の会話を分析した研究では、選択された発話者が応答を遅延させると、他の非選択発話者が応答を試みる傾向が強くなることが明らかになっている [73]。また、H. Maclay と C. E. Osgood らは、イリノイ大学で開催されたカンファレンスから採取された英語の発話データを分析し、スピーチ内のためらいには 4 つの種類が存在することを示している [74]。加えて、話者のためらいは聞き手の理解にも影響を与えることが指摘されている [75]。このような機微を、先述した拡張で取り扱うのは困難だと言える。また、組織内における沈黙が、集合的な現象であることを示唆する複数の研究があるが [70, 72, 76–81]、これらは、従業員が組織のヒエラルキーを意識したり、あるいは、声を上げることを無駄だと感じた結果であったりするため、同じく本モデルで扱えるような単純なものではない。これらは性差や文化的要因の差によっても引き起こされる可能性がある点も重要である [81]。

4 結論

本論文では、集団における沈黙という現象を、沈黙 (0) と発話 (1) の 2 値状態を取るエージェントのマルコフ連鎖モデルによって解析した。まずは互いに相互作用のない孤立集団を考え、次に均一な相互作用の場合を考察し、最後にそれらを補間する形で喫茶店 θ モデルを導出した。多くの近似を用いたにもかかわらず、喫茶店 θ モデルとシミュレーションの一致は良好で、重みをランダムにした場合でも、このモデルで分布の大まかな傾向を捉えられることが分かった。これは、このモデルの幅広い説明可能性を示唆する結果である。しかし、重みがランダムで、かつ、場の人数 N が少ないときに発生する多峰性は、単純なベータ分布で記述される喫茶店 θ モデルでは説明できなかった。これは、喫茶店 θ モデルの限界を示すとともに、今回のマルコフ連鎖モデル自体には、より多くの分布を表現する能力があることを意味する。

また、この喫茶店 θ モデルを利用して沈黙の発生条件を調べたところ、各個人が現状を維持する傾向が強い、あるいは場の人数が少ないときに沈黙が発生しやすいという結果が得られた。これは我々の直感とも一致する。一方で、周囲の会話量に対する各々の意識を下げていくと、あるところで沈黙の発生確率が急激に 0 になる、相転移のような非自明の現象が見られた。フランスのことわざを借りれば、天使が通るのは、各々がわずかにでも周囲に意識を向けたときだけだと言うことができるだろう。

さらに我々は、4 人のディナーパーティの様子を収録した CHiME-6 データセットと喫茶店 θ モデルの比較も試みた。喫茶店 θ モデルは、大人数に対するモデルであるため、この比較が適切ではないことは再度強調しておきたい。しかし、喫茶店 θ モデルでも、その分布をある程度再現可能であることが明らかになった。ただし、少人数におけるシミュレーションのときと同様に、この音声データにも多峰性が見られ、その解析のため

にはさらなる調査が必要である。

最後に、実際の会話の持つ特徴について調査し、提案モデルの課題と限界について議論した。

今回は主に、会話量の分布という静的な量に着目して解析を行った。しかし、このモデルはダイナミクスを調べる上でも有用である。動的な人間のコミュニケーションの解析は、複雑性科学の文脈においても重要視されている [82]。今後は、本モデルを通して沈黙前後のダイナミクスを調べ、その構造を明らかにすることが期待される。加えて、実験値との比較はまだ不十分である。特に、喫茶店 θ モデルの正確な検証を行うためには、十分な人数が同じ空間にいる場合の自然な会話データが必要である。そして、実験データと我々のモデルをどのように比較すべきか、つまり、 α, β, θ などのパラメータを現実とどう対応させるべきかということについても議論が必要である。例えば、M. Cristani らの行った解析では、様々な対話シナリオにおける発話と沈黙の時間の分布が明らかになっており [42]、それに基づいた様々な α, β を用意することで、実際の会話を反映した沈黙の発生条件が調べられる可能性がある。あるいは、平均沈黙時間や、平均発話時間といった実験で計測可能な量で α, β を推定し、会話量の分布が実験値に合うように θ を調整することで、直接測定することが困難な他人の会話に対する意識 θ を評価できるようになるかもしれない。

5 Code Availability

The source code for the model discussed in this study is publicly accessible on GitHub <https://github.com/A5size/angel-passes> and is licensed under the terms of the MIT License.

6 Acknowledgments

I would like to thank Kentaro Yonemura for his valuable feedback. I am also deeply grateful to Hirokazu Nitta for his generous support and encouragement.

付録 A 均一集団分布の導出

$w_{ij} = 1/N$ の場合に、十分時間が経った後で n_1 人が話している確率 $f_{\text{uni}}(n_1)$ が満たすべき方程式は、

$$f_{\text{uni}}(n'_1) = \sum_{n_1=0}^N g(n_1, n'_1) f_{\text{uni}}(n_1)$$

となる。ここで

$$g(n_1, n'_1) = \binom{N}{n'_1} \left(\frac{n_0}{N} \alpha + \frac{n_1}{N} (1 - \beta) \right)^{n'_1} \left(\frac{n_0}{N} (1 - \alpha) + \frac{n_1}{N} \beta \right)^{N - n'_1}$$

である。

一般にこれを解くことは難しい。そこで、まずはこれを積分方程式に書き直し、パラメータを含む形で解を仮定した後で、尤もらしいパラメータを決定することで解を近似する。

まずは、

$$x = \frac{n_1}{N}$$

$$x' = \frac{n'_1}{N}$$

において、先の方程式を

$$\begin{aligned} \tilde{f}_{\text{uni}}(x') &= \frac{\Gamma(N+1)}{\Gamma(x'N+1)\Gamma(N-x'N+1)} \\ &\quad \times N \int_0^1 \{(1-x)\alpha + x(1-\beta)\}^{x'N} \{(1-x)(1-\alpha) + x\beta\}^{N-x'N} \tilde{f}_{\text{uni}}(x) dx \end{aligned}$$

という積分方程式に書き換える。

ここで

$$y = (1-x)\alpha + x(1-\beta)$$

$$y' = (1-x')\alpha + x'(1-\beta)$$

とすると、先の積分は

$$\tilde{f}_{\text{uni}}(y') = \frac{\Gamma(N+1)}{\Gamma(x'N+1)\Gamma(N-x'N+1)} \frac{N}{1-\alpha-\beta} \int_{\alpha}^{1-\beta} y^{x'N} (1-y)^{N-x'N} \tilde{f}_{\text{uni}}(y) dy$$

となる。

$\tilde{f}_{\text{uni}}(x)$ を、未知のパラメータ a, b とベータ関数 B を使って

$$\tilde{f}(x) = \frac{1}{B(a, b)} y^{a-1} (1-y)^{b-1}$$

と書けると仮定すると

$$y'^{a-1} (1-y')^{b-1} = \frac{\Gamma(N+1)}{\Gamma(x'N+1)\Gamma(N-x'N+1)} \frac{N}{1-\alpha-\beta} \int_{\alpha}^{1-\beta} y^{a+x'N-1} (1-y)^{b+N-x'N-1} dy$$

となり、 $0 < \alpha, \beta \ll 1$ であるので、積分区間を 0 から 1 と近似すると、この積分は

$$\begin{aligned} &\int_{\alpha}^{1-\beta} y^{a+x'N-1} (1-y)^{b+N-x'N-1} dy \\ &\sim \int_0^1 y^{a+x'N-1} (1-y)^{b+N-x'N-1} dy = \frac{\Gamma(a+x'N)\Gamma(b+N-x'N)}{\Gamma(a+b+N)} \end{aligned}$$

と計算できる。

これより、パラメータ a, b を決定するための方程式として

$$y'^{a-1} (1-y')^{b-1} = \frac{\Gamma(N+1)}{\Gamma(x'N+1)\Gamma(N-x'N+1)} \frac{N}{1-\alpha-\beta} \frac{\Gamma(a+x'N)\Gamma(b+N-x'N)}{\Gamma(a+b+N)}$$

を得る。

ここでは、パラメータ a, b を、 $x' = 1/2$ のときに両辺がほとんど一致するように決定する。これは、スターリングの公式から、 z が十分大きいとき $\Gamma(z)$ を

$$\Gamma(z) \sim \sqrt{\frac{2\pi}{z}} \left(\frac{z}{e}\right)^z$$

と近似できるためである。

まずガンマ関数を置き換えると

$$\begin{aligned} & \{(1-x')\alpha + x'(1-\beta)\}^{a-1} \{(1-x')(1-\alpha) + x'\beta\}^{b-1} \\ &= \frac{\sqrt{N}}{1-\alpha-\beta} \frac{(a+x'N)^{a+x'N-\frac{1}{2}} (b+N-x'N)^{b+N-x'N-\frac{1}{2}}}{x'^{x'N+\frac{1}{2}} (1-x')^{N-x'N+\frac{1}{2}} (a+b+N)^{a+b+N-\frac{1}{2}}} \end{aligned}$$

となり、 $x' = 1/2$ を代入すると

$$(1+\alpha-\beta)^{a-1} (1+\beta-\alpha)^{b-1} = \frac{1}{1-\alpha-\beta} \frac{\left(\frac{2a}{N}+1\right)^{a+\frac{N-1}{2}} \left(\frac{2b}{N}+1\right)^{b+\frac{N-1}{2}}}{\left(\frac{a+b}{N}+1\right)^{a+b+N-\frac{1}{2}}}$$

を得る。

ここで、新たなパラメータ c を使って、 $a = c\alpha N, b = c\beta N$ とおきなおすと、

$$(1+\alpha-\beta)^{c\alpha N-1} (1+\beta-\alpha)^{c\beta N-1} = \frac{1}{1-\alpha-\beta} \frac{(2c\alpha+1)^{c\alpha N+\frac{N-1}{2}} (2c\beta+1)^{c\beta N+\frac{N-1}{2}}}{(c\alpha+c\beta+1)^{c\alpha N+c\beta N+N-\frac{1}{2}}}$$

となる。

これを c について解くために、両辺に \log をとり、 $0 < \alpha, \beta \ll 1$ であることを使って、 α, β の 2 次以上の項を落とし、 $\log(1+x) \sim x$ を使えば、最終的に $c = 2$ が得られる。

以上より、 $\tilde{f}_{\text{uni}}(x)$ は、

$$\tilde{f}_{\text{uni}}(x) = \frac{1}{B(2\alpha N, 2\beta N)} \{(1-x)\alpha + x(1-\beta)\}^{2\alpha N-1} \{(1-x)(1-\alpha) + x\beta\}^{2\beta N-1}$$

となる。

この導出において、 $x = n_1/N$ を連続量として扱ってしまっているため、この表式を、そのまま本来の離散的量に適用してしまうと、全事象の確率の和が 1 ではなくなってしまう。そこで、離散的な場合には、規格化の定数を新たにおきなとして

$$f_{\text{uni}}(n_1) = \frac{1}{Z(2\alpha N, 2\beta N)} \left\{ \left(1 - \frac{n_1}{N}\right) \alpha + \frac{n_1}{N} (1 - \beta) \right\}^{2\alpha N - 1} \left\{ \left(1 - \frac{n_1}{N}\right) (1 - \alpha) + \frac{n_1}{N} \beta \right\}^{2\beta N - 1}$$

とする必要がある。

ここで、

$$Z(2\alpha N, 2\beta N) = \sum_{n_1=0}^N \left\{ \left(1 - \frac{n_1}{N}\right) \alpha + \frac{n_1}{N} (1 - \beta) \right\}^{2\alpha N - 1} \left\{ \left(1 - \frac{n_1}{N}\right) (1 - \alpha) + \frac{n_1}{N} \beta \right\}^{2\beta N - 1}$$

であり、 N が十分大きい場合には

$$Z(2\alpha N, 2\beta N) \sim \frac{N}{1 - \alpha - \beta} B(2\alpha N, 2\beta N)$$

と近似できる。

また、それほど N が大きくない場合でも、台形法の考え方から

$$Z(2\alpha N, 2\beta N) \sim \frac{N}{1 - \alpha - \beta} B(2\alpha N, 2\beta N) + \frac{1}{2} (\alpha^{2\alpha N - 1} (1 - \alpha)^{2\beta N - 1} + (1 - \beta)^{2\alpha N - 1} \beta^{2\beta N - 1})$$

と近似できる。

参考文献

- [1] Maurizio Bettini and William Michael Short. *Hermes' Ears: Places and Symbols of Communication in ancient Culture*, pages 3–39. Ohio State University Press, 2000. ISBN 9780814211700. URL <http://www.jstor.org/stable/j.ctv16b7842.5>.
- [2] Deborah Tannen. Silence: anything but. *Perspectives on silence*, 93:93–111, 1985.
- [3] Philip K. Bock, Deborah Tannen, and Muriel Saville-Troike. Perspectives on silence. *Language*, 62(3):731, September 1986. ISSN 0097-8507. doi:[10.2307/415529](https://doi.org/10.2307/415529). URL <http://dx.doi.org/10.2307/415529>.
- [4] Ikuko Nakane. Silence, January 2012. URL <http://dx.doi.org/10.1002/9781118247273.ch9>.
- [5] Adam Jaworski. *The Power of Silence: Social and Pragmatic Perspectives*. SAGE Publications, Inc., 1993. ISBN 9781483325460. doi:[10.4135/9781483325460](https://doi.org/10.4135/9781483325460). URL <http://dx.doi.org/10.4135/9781483325460>.

- [6] Sally Jones. Speech is silver, silence is golden: The cultural importance of silence in japan. *ANU Undergraduate Research Journal*, 3, December 2011. ISSN 1837-2872. doi:[10.22459/aurj.03.2011.02](https://doi.org/10.22459/aurj.03.2011.02). URL <http://dx.doi.org/10.22459/aurj.03.2011.02>.
- [7] Daniel Steinbock. *Making silence together: Collaboration in the silent gatherings of Quakers*. Stanford University, 2012.
- [8] Elizabeth Molina-Markham. Finding the “sense of the meeting” : Decision making through silence among quakers. *Western Journal of Communication*, 78(2):155–174, 2014. doi:[10.1080/10570314.2013.809474](https://doi.org/10.1080/10570314.2013.809474). URL <https://doi.org/10.1080/10570314.2013.809474>.
- [9] Jeff Waistell. The salience of silence: the silence of salience. *Journal of Management, Spirituality & Religion*, 15(3):211–228, May 2018. ISSN 1476-6086. doi:[10.1080/14766086.2018.1449136](https://doi.org/10.1080/14766086.2018.1449136). URL <http://dx.doi.org/10.1080/14766086.2018.1449136>.
- [10] Zhixu (Rick) Yang and Franki Y. H. Kung. Toward a culturally sensitive perspective on silence in organizations. *Industrial and Organizational Psychology*, 17(3):366–370, September 2024. ISSN 1754-9434. doi:[10.1017/iop.2024.24](https://doi.org/10.1017/iop.2024.24). URL <http://dx.doi.org/10.1017/iop.2024.24>.
- [11] Keith H. Basso. “to give up on words”: Silence in western apache culture. *Southwestern Journal of Anthropology*, 26(3):213–230, 1970. doi:[10.1086/soutjanth.26.3.3629378](https://doi.org/10.1086/soutjanth.26.3.3629378). URL <https://doi.org/10.1086/soutjanth.26.3.3629378>.
- [12] Philip M. Peek. The sounds of silence: Cross-world communication and the auditory arts in african societies. *American Ethnologist*, 21(3):474–494, 1994. ISSN 00940496, 15481425. URL <http://www.jstor.org/stable/645917>.
- [13] Kofi Agyekum. The communicative role of silence in akan. *Pragmatics*, page 31–51, July 2022. ISSN 2406-4238. doi:[10.1075/prag.12.1.03agy](https://doi.org/10.1075/prag.12.1.03agy). URL <http://dx.doi.org/10.1075/prag.12.1.03agy>.
- [14] Adam Jaworski and Itesh Sachdev. Beliefs about silence in the classroom. *Language and Education*, 12(4):273–292, December 1998. ISSN 1747-7581. doi:[10.1080/09500789808666754](https://doi.org/10.1080/09500789808666754). URL <http://dx.doi.org/10.1080/09500789808666754>.
- [15] Dennis Kurzon. On silence. *Journal of Pragmatics*, 43(9):2275–2277, July 2011. ISSN 0378-2166. doi:[10.1016/j.pragma.2010.11.011](https://doi.org/10.1016/j.pragma.2010.11.011). URL <http://dx.doi.org/10.1016/j.pragma.2010.11.011>.
- [16] Jared R. Curhan, Jennifer R. Overbeck, Yeri Cho, Teng Zhang, and Yu Yang. Silence is golden: Extended silence, deliberative mindset, and value creation in negotiation. *Journal of Applied Psychology*, 107(1):78–94, January 2022. ISSN 0021-9010. doi:[10.1037/apl0000877](https://doi.org/10.1037/apl0000877). URL <http://dx.doi.org/10.1037/apl0000877>.
- [17] 種市瑛. 会話に見られる沈黙の解釈の多義性—語用実践行為として捉える沈黙—. 社会言語科学会第 43 回大会発表論文集, pages 122–125, 3 2019.
- [18] Amy C. Edmondson and Tijs Besieux. Reflections: Voice and silence in workplace conversations. *Journal of Change Management*, 21(3):269–286, May 2021. ISSN 1479-1811. doi:[10.1080/14697017.2021.1928910](https://doi.org/10.1080/14697017.2021.1928910). URL <http://dx.doi.org/10.1080/14697017.2021.1928910>.
- [19] Eric Pfeifer and Marc Wittmann. Waiting, thinking, and feeling: Variations in the perception of time during silence. *Frontiers in Psychology*, 11, 2020. ISSN 1664-1078. doi:[10.3389/fpsyg.2020.00602](https://doi.org/10.3389/fpsyg.2020.00602). URL <https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2020.00602>.

- [20] Alexandre Asselineau, Gilles Grolleau, and Naoufel Mzoughi. Quiet environments and the intentional practice of silence: Toward a new perspective in the analysis of silence in organizations. *Industrial and Organizational Psychology*, 17(3):326-340, May 2024. ISSN 1754-9434. doi:[10.1017/iop.2024.9](https://doi.org/10.1017/iop.2024.9). URL <http://dx.doi.org/10.1017/iop.2024.9>.
- [21] Lesley Saunders. Silences and silence in “creativity” . *London Review of Education*, 10(2), 2012. ISSN 1474-8479. doi:[10.1080/14748460.2012.691285](https://doi.org/10.1080/14748460.2012.691285). URL <http://dx.doi.org/10.1080/14748460.2012.691285>.
- [22] Vinca Bigo. On silence, creativity and ethics in organization studies. *Organization Studies*, 39(1): 121-133, August 2017. ISSN 1741-3044. doi:[10.1177/0170840617717553](https://doi.org/10.1177/0170840617717553). URL <http://dx.doi.org/10.1177/0170840617717553>.
- [23] 真友里橋本. 心理臨床における沈黙の概念および沈黙研究の概観と展望. 京都大学大学院教育学研究科紀要, 62, 3 2016. URL <http://hdl.handle.net/2433/209908>.
- [24] Erica C. Kaye, Sarah L. Rockwell, Monica E. Lemmon, Justin N. Baker, and Jennifer W. Mack. The art of saying nothing. *Pediatrics*, 149(6), May 2022. ISSN 1098-4275. doi:[10.1542/peds.2022-056862](https://doi.org/10.1542/peds.2022-056862). URL <http://dx.doi.org/10.1542/peds.2022-056862>.
- [25] Anthony L. Back, Susan M. Bauer-Wu, Cynda H. Rushton, and Joan Halifax. Compassionate silence in the patient-clinician encounter: A contemplative approach. *Journal of Palliative Medicine*, 12 (12):1113-1117, December 2009. ISSN 1557-7740. doi:[10.1089/jpm.2009.0175](https://doi.org/10.1089/jpm.2009.0175). URL <http://dx.doi.org/10.1089/jpm.2009.0175>.
- [26] 剛志糸林. カウンセリングにおける沈黙の意味について. 人間性心理学研究, 27(1-2):57-68, 2009. ISSN 02894904. URL <https://cir.nii.ac.jp/crid/1050865884350959360>.
- [27] Josef Bartels, Rachel Rodenbach, Katherine Ciesinski, Robert Gramling, Kevin Fiscella, and Ronald Epstein. Eloquent silences: A musical and lexical analysis of conversation between oncologists and their patients. *Patient Education and Counseling*, 99(10):1584-1594, October 2016. ISSN 0738-3991. doi:[10.1016/j.pec.2016.04.009](https://doi.org/10.1016/j.pec.2016.04.009). URL <http://dx.doi.org/10.1016/j.pec.2016.04.009>.
- [28] H.B. Kekre, C.L. Saxena, and H.M. Srivastava. A two-state markov model of speech in conversation and its application to computer communication systems. *Computers & Electrical Engineering*, 4(2): 133-141, 1977. ISSN 0045-7906. doi:[https://doi.org/10.1016/0045-7906\(77\)90022-2](https://doi.org/10.1016/0045-7906(77)90022-2). URL <https://www.sciencedirect.com/science/article/pii/0045790677900222>.
- [29] OI Sheluhin, AA Atayero, MO Eyinagho, and JO Iruemi. Markov models of telephone speech dialogue. *International Journal of Engineering and Technology*, 2(3):328-333, 2012.
- [30] Michel Galley, Kathleen McKeown, Eric Fosler-Lussier, and Hongyan Jing. Discourse segmentation of multi-party conversation. In *Proceedings of the 41st Annual Meeting on Association for Computational Linguistics - ACL ' 03*, volume 1 of *ACL ' 03*, page 562-569. Association for Computational Linguistics, 2003. doi:[10.3115/1075096.1075167](https://doi.org/10.3115/1075096.1075167). URL <http://dx.doi.org/10.3115/1075096.1075167>.
- [31] Elliott M. Hoey and Kobin H. Kendrick. Conversation analysis, October 2017. URL <http://dx.doi.org/10.1002/9781394259762.ch8>.
- [32] Andrew P. Thomas, Peter Bull, and Derek Roger. Conversational exchange analysis. *Journal of Language and Social Psychology*, 1(2):141-155, September 1982. ISSN 1552-6526.

- doi:10.1177/0261927x8200100204. URL <http://dx.doi.org/10.1177/0261927x8200100204>.
- [33] Masato Ishizaki and Tsuneaki Kato. Exploring the characteristics of multi-party dialogues. In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics - Volume 1*, ACL '98/COLING '98, page 583-589, USA, 1998. Association for Computational Linguistics. doi:10.3115/980845.980942. URL <https://doi.org/10.3115/980845.980942>.
 - [34] Andreas Stolcke, Klaus Ries, Noah Coccaro, Elizabeth Shriberg, Rebecca Bates, Daniel Jurafsky, Paul Taylor, Rachel Martin, Carol Van Ess-Dykema, and Marie Meteer. Dialogue act modeling for automatic tagging and recognition of conversational speech. *Computational Linguistics*, 26(3): 339-374, 2000. URL <https://aclanthology.org/J00-3003/>.
 - [35] Paul M. Aoki, Matthew Romaine, Margaret H. Szymanski, James D. Thornton, Daniel Wilson, and Allison Woodruff. The mad hatter's cocktail party: a social mobile audio space supporting multiple simultaneous conversations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '03, page 425-432, New York, NY, USA, 2003. Association for Computing Machinery. ISBN 1581136307. doi:10.1145/642611.642686. URL <https://doi.org/10.1145/642611.642686>.
 - [36] Sumit Basu, Tanzeem Choudhury, and Brian Clarkson. Learning human interactions with the influence model. Technical Report MIT-TR-539, June 2001. URL <https://www.microsoft.com/en-us/research/publication/learning-human-interactions-influence-model/>. MIT Media Laboratory Vision and Modeling Technical Report #539.
 - [37] Wen Dong, Bruno Lepri, Alessandro Cappelletti, Alex Sandy Pentland, Fabio Pianesi, and Massimo Zancanaro. Using the influence model to recognize functional roles in meetings. In *Proceedings of the 9th International Conference on Multimodal Interfaces*, ICMI '07, page 271-278, New York, NY, USA, 2007. Association for Computing Machinery. ISBN 9781595938176. doi:10.1145/1322192.1322239. URL <https://doi.org/10.1145/1322192.1322239>.
 - [38] Kazuhiro Otsuka, Yoshinao Takemae, and Junji Yamato. A probabilistic inference of multiparty-conversation structure based on markov-switching models of gaze patterns, head directions, and utterances. In *Proceedings of the 7th International Conference on Multimodal Interfaces*, ICMI '05, page 191-198, New York, NY, USA, 2005. Association for Computing Machinery. ISBN 1595930280. doi:10.1145/1088463.1088497. URL <https://doi.org/10.1145/1088463.1088497>.
 - [39] Manabu Fujimoto and Ikuo Daibo. Conversation development patterns in small group interactions. *THE JAPANESE JOURNAL OF EXPERIMENTAL SOCIAL PSYCHOLOGY*, 47(1):51-60, 2007. ISSN 1348-6276. doi:10.2130/jjesp.47.51. URL <http://dx.doi.org/10.2130/jjesp.47.51>.
 - [40] Tony Bergstrom and Karrie Karahalios. *Seeing More: Visualizing Audio Cues*, page 29-42. Springer Berlin Heidelberg, 2007. ISBN 9783540748007. doi:10.1007/978-3-540-74800-7_3. URL http://dx.doi.org/10.1007/978-3-540-74800-7_3.
 - [41] Massimo Mastrangeli, Martin Schmidt, and Lucas Lacasa. The roundtable: an abstract model of conversation dynamics, 2010. URL <https://arxiv.org/abs/1010.2943>.
 - [42] Marco Cristani, Anna Pesarin, Carlo Drioli, Alessandro Tavano, Alessandro Perina, and Vittorio Murino. Generative modeling and classification of dialogs by a low-level turn-taking feature. *Pattern*

- Recogn.*, 44(8):1785-1800, August 2011. ISSN 0031-3203. doi:[10.1016/j.patcog.2011.01.013](https://doi.org/10.1016/j.patcog.2011.01.013). URL <https://doi.org/10.1016/j.patcog.2011.01.013>.
- [43] D. Angus, A. Smith, and J. Wiles. Conceptual recurrence plots: Revealing patterns in human discourse. *IEEE Transactions on Visualization and Computer Graphics*, 18(6):988-997, June 2012. ISSN 1077-2626. doi:[10.1109/tvcg.2011.100](https://doi.org/10.1109/tvcg.2011.100). URL <http://dx.doi.org/10.1109/TVCG.2011.100>.
- [44] Daniel Angus, Bernadette Watson, Andrew Smith, Cindy Gallois, and Janet Wiles. Visualising conversation structure across time: Insights into effective doctor-patient consultations. *PLoS ONE*, 7(6):e38014, June 2012. ISSN 1932-6203. doi:[10.1371/journal.pone.0038014](https://doi.org/10.1371/journal.pone.0038014). URL <http://dx.doi.org/10.1371/journal.pone.0038014>.
- [45] Riccardo Fusaroli and Kristian Tylén. Investigating conversational dynamics: Interactive alignment, interpersonal synergy, and collective task performance. *Cognitive Science*, 40(1):145-171, May 2015. ISSN 1551-6709. doi:[10.1111/cogs.12251](https://doi.org/10.1111/cogs.12251). URL <http://dx.doi.org/10.1111/cogs.12251>.
- [46] Laurence A. Clarfeld, Robert Gramling, Donna M. Rizzo, and Margaret J. Eppstein. A general model of conversational dynamics and an example application in serious illness communication. *PLOS ONE*, 16(7):e0253124, July 2021. ISSN 1932-6203. doi:[10.1371/journal.pone.0253124](https://doi.org/10.1371/journal.pone.0253124). URL <http://dx.doi.org/10.1371/journal.pone.0253124>.
- [47] Larry S. Liebovitch, Paul R. Peluso, Michael D. Norman, Jessica Su, and John M. Gottman. Mathematical model of the dynamics of psychotherapy. *Cognitive Neurodynamics*, 5(3):265-275, May 2011. ISSN 1871-4099. doi:[10.1007/s11571-011-9157-x](https://doi.org/10.1007/s11571-011-9157-x). URL <http://dx.doi.org/10.1007/s11571-011-9157-x>.
- [48] Georg Simmel. The number of members as determining the sociological form of the group. i. *American Journal of Sociology*, 8(1):1-46, July 1902. ISSN 1537-5390. doi:[10.1086/211115](https://doi.org/10.1086/211115). URL <http://dx.doi.org/10.1086/211115>.
- [49] Aldha Willian. Social contexts and conversational implicatures in conversations among family members. *ELT Echo: The Journal of English Language Teaching in Foreign Language Context*, 3(2):169, December 2018. ISSN 2579-8170. doi:[10.24235/eltecho.v3i2.3244](https://doi.org/10.24235/eltecho.v3i2.3244). URL <http://dx.doi.org/10.24235/eltecho.v3i2.3244>.
- [50] Miles Cranmer. Interpretable machine learning for science with pysr and symbolicregression.jl, 2023. URL <https://arxiv.org/abs/2305.01582>.
- [51] Shinji Watanabe, Michael Mandel, Jon Barker, Emmanuel Vincent, Ashish Arora, Xuankai Chang, Sanjeev Khudanpur, Vimal Manohar, Daniel Povey, Desh Raj, David Snyder, Aswin Shanmugam Subramanian, Jan Trmal, Bar Ben Yair, Christoph Boeddeker, Zhaoheng Ni, Yusuke Fujita, Shota Horiguchi, Naoyuki Kanda, Takuya Yoshioka, and Neville Ryant. Chime-6 challenge: Tackling multispeaker speech recognition for unsegmented recordings. In *6th International Workshop on Speech Processing in Everyday Environments (CHiME 2020)*. ISCA, May 2020. doi:[10.21437/chime.2020-1](https://doi.org/10.21437/chime.2020-1). URL <http://dx.doi.org/10.21437/CHIME.2020-1>.
- [52] SCHEGLOFF E. A. Discourse as an interactional achievement : Some uses of 'uh huh' and other things that come between sentences. *Analyzing Discourse : Text and Talk (Georgetown University Roundtable on Languages and Linguistics 1981)*, pages 71-93, 1982. URL <https://cir.nii.ac.jp/crid/1571135650710026880>.

- [53] Chirag Raman and Hayley Hung. Towards automatic estimation of conversation floors within f-formations. In *2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*, page 175-181. IEEE, September 2019. doi:[10.1109/aciw.2019.8925065](https://doi.org/10.1109/aciw.2019.8925065). URL <http://dx.doi.org/10.1109/ACIIW.2019.8925065>.
- [54] Paul M. Aoki, Margaret H. Szymanski, Luke Plurkowski, James D. Thornton, Allison Woodruff, and Weilie Yi. Where’s the “party” in “multi-party”? analyzing the structure of small-group sociable talk. In *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work, CSCW '06*, page 393-402, New York, NY, USA, 2006. Association for Computing Machinery. ISBN 1595932496. doi:[10.1145/1180875.1180934](https://doi.org/10.1145/1180875.1180934). URL <https://doi.org/10.1145/1180875.1180934>.
- [55] Edward Twitchell Hall and 三郎大江. *The silent language*. Nan’un-do’s contemporary library. Nan’un-do, 1966. URL <https://cir.nii.ac.jp/crid/1130000794457039104>.
- [56] J. L. (John Langshaw) Austin, J. O. (James Opie) Urmson, and Marina Sbisà. *How to do things with words*. Number 1955 in William James lectures. Oxford University Press, 2nd ed. / edited by j.o. urmson and marina sbisà edition, 1975. URL <https://ci.nii.ac.jp/ncid/BA01092074>.
- [57] Jürgen Streeck. The significance of gesture: How it is established. *IPrA Papers in Pragmatics*, page 60-83, July 2022. ISSN 2406-4246. doi:[10.1075/iprapip.2.1-2.03str](https://doi.org/10.1075/iprapip.2.1-2.03str). URL <http://dx.doi.org/10.1075/iprapip.2.1-2.03str>.
- [58] Cecilia E. Ford, Barbara A. Fox, and Sandra A. Thompson. Practices in the construction of turns: The “tcu” revisited. *Pragmatics*, page 427-454, July 2022. ISSN 2406-4238. doi:[10.1075/prag.6.3.07for](https://doi.org/10.1075/prag.6.3.07for). URL <http://dx.doi.org/10.1075/prag.6.3.07for>.
- [59] Herbert H. Clark. *Using Language*. Cambridge University Press, May 1996. ISBN 9780511620539. doi:[10.1017/cbo9780511620539](https://doi.org/10.1017/cbo9780511620539). URL <http://dx.doi.org/10.1017/CB09780511620539>.
- [60] Mayumi Bono, Noriko Suzuki, and Yasuhiro Katagiri. An analysis of participation structure in conversation based on interaction corpus of ubiquitous sensor data. In *IFIP TC13 International Conference on Human-Computer Interaction*, 2003. URL <https://api.semanticscholar.org/CorpusID:1167126>.
- [61] Roel Vertegaal, Ivo Weevers, Changuk Sohn, and Chris Cheung. Gaze-2: conveying eye contact in group video conferencing using eye-controlled camera direction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '03*, page 521-528, New York, NY, USA, 2003. Association for Computing Machinery. ISBN 1581136307. doi:[10.1145/642611.642702](https://doi.org/10.1145/642611.642702). URL <https://doi.org/10.1145/642611.642702>.
- [62] Alex Pentland. Social dynamics: Signals and behavior. In *International Conference on Developmental Learning*, volume 5, pages 1-5, 2004.
- [63] Tanya Stivers. Stance, alignment, and affiliation during storytelling: When nodding is a token of affiliation. *Research on Language and Social Interaction*, 41(1):31-57, February 2008. ISSN 1532-7973. doi:[10.1080/08351810701691123](https://doi.org/10.1080/08351810701691123). URL <http://dx.doi.org/10.1080/08351810701691123>.
- [64] Tanya Stivers, N. J. Enfield, Penelope Brown, Christina Englert, Makoto Hayashi, Trine Heinemann, Gertie Hoymann, Federico Rossano, Jan Peter de Ruiter, Kyung-Eun Yoon, and Stephen C. Levinson. Universals and cultural variation in turn-taking in conversation. *Proceed-*

- ings of the National Academy of Sciences*, 106(26):10587-10592, June 2009. ISSN 1091-6490. doi:[10.1073/pnas.0903616106](https://doi.org/10.1073/pnas.0903616106). URL <http://dx.doi.org/10.1073/pnas.0903616106>.
- [65] Kobin H. Kendrick, Judith Holler, and Stephen C. Levinson. Turn-taking in human face-to-face interaction is multimodal: gaze direction and manual gestures aid the coordination of turn transitions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 378(1875), March 2023. ISSN 1471-2970. doi:[10.1098/rstb.2021.0473](https://doi.org/10.1098/rstb.2021.0473). URL <http://dx.doi.org/10.1098/rstb.2021.0473>.
- [66] 知己石坂, バックホー, and 大妹尾. 二者のオンライン会話における沈黙研究. 経営情報学会全国研究発表大会要旨集, 202111(0):59-62, 01 2022. doi:[10.11497/jasmin.202111.0_59](https://doi.org/10.11497/jasmin.202111.0_59). URL <https://cir.nii.ac.jp/crid/1390009393790152192>.
- [67] EMANUEL A. SCHEGLOFF. Overlapping talk and the organization of turn-taking for conversation. *Language in Society*, 29(1):1-63, January 2000. ISSN 1469-8013. doi:[10.1017/S0047404500001019](https://doi.org/10.1017/S0047404500001019). URL <http://dx.doi.org/10.1017/S0047404500001019>.
- [68] Dennis Kurzon. The right of silence: A socio-pragmatic model of interpretation. *Journal of Pragmatics*, 23(1):55-69, 1995. ISSN 0378-2166. doi:[https://doi.org/10.1016/0378-2166\(94\)00036-E](https://doi.org/10.1016/0378-2166(94)00036-E). URL <https://www.sciencedirect.com/science/article/pii/037821669400036E>. Laying down the law - Discourse analysis of legal institutions.
- [69] Dennis Kurzon. Towards a typology of silence. *Journal of Pragmatics*, 39(10):1673-1688, 2007. ISSN 0378-2166. doi:<https://doi.org/10.1016/j.pragma.2007.07.003>. URL <https://www.sciencedirect.com/science/article/pii/S037821660700121X>.
- [70] Chad T. Brinsfield. Employee silence motives: Investigation of dimensionality and development of measures. *Journal of Organizational Behavior*, 34(5):671-697, 2013. doi:<https://doi.org/10.1002/job.1829>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/job.1829>.
- [71] Charles B. Stone, Alin Coman, Adam D. Brown, Jonathan Koppel, and William Hirst. Toward a science of silence: The consequences of leaving a memory unsaid. *Perspectives on Psychological Science*, 7(1):39-53, January 2012. ISSN 1745-6924. doi:[10.1177/1745691611427303](https://doi.org/10.1177/1745691611427303). URL <http://dx.doi.org/10.1177/1745691611427303>.
- [72] Leilei Hao, Hui Zhu, Yuqian He, Jinyun Duan, Teng Zhao, and Hui Meng. When is silence golden? a meta-analysis on antecedents and outcomes of employee silence. *Journal of Business and Psychology*, 37(5):1039-1063, February 2022. ISSN 1573-353X. doi:[10.1007/s10869-021-09788-7](https://doi.org/10.1007/s10869-021-09788-7). URL <http://dx.doi.org/10.1007/s10869-021-09788-7>.
- [73] TANYA STIVERS and JEFFREY D. ROBINSON. A preference for progressivity in interaction. *Language in Society*, 35(03), May 2006. ISSN 1469-8013. doi:[10.1017/S0047404506060179](https://doi.org/10.1017/S0047404506060179). URL <http://dx.doi.org/10.1017/S0047404506060179>.
- [74] Howard Maclay and Charles E. Osgood. Hesitation phenomena in spontaneous english speech. *WORD*, 15(1):19-44, January 1959. ISSN 2373-5112. doi:[10.1080/00437956.1959.11659682](https://doi.org/10.1080/00437956.1959.11659682). URL <http://dx.doi.org/10.1080/00437956.1959.11659682>.
- [75] James G. Martin and Winifred Strange. The perception of hesitation in spontaneous speech. *Perception & Psychophysics*, 3(6):427-438, November 1968. ISSN 1532-5962. doi:[10.3758/bf03205750](https://doi.org/10.3758/bf03205750). URL <http://dx.doi.org/10.3758/BF03205750>.

- [76] Frances J. Milliken, Elizabeth W. Morrison, and Patricia F. Hewlin. An exploratory study of employee silence: Issues that employees don't communicate upward and why*. *Journal of Management Studies*, 40(6):1453-1476, August 2003. ISSN 1467-6486. doi:[10.1111/1467-6486.00387](https://doi.org/10.1111/1467-6486.00387). URL <http://dx.doi.org/10.1111/1467-6486.00387>.
- [77] Jimmy Donaghey, Niall Cullinane, Tony Dundon, and Adrian Wilkinson. Reconceptualising employee silence: problems and prognosis. *Work, Employment and Society*, 25(1):51-67, March 2011. ISSN 1469-8722. doi:[10.1177/0950017010389239](https://doi.org/10.1177/0950017010389239). URL <http://dx.doi.org/10.1177/0950017010389239>.
- [78] Elizabeth W. Morrison. Employee voice and silence. *Annual Review of Organizational Psychology and Organizational Behavior*, 1(1):173-197, March 2014. ISSN 2327-0616. doi:[10.1146/annurev-orgpsych-031413-091328](https://doi.org/10.1146/annurev-orgpsych-031413-091328). URL <http://dx.doi.org/10.1146/annurev-orgpsych-031413-091328>.
- [79] Michael Knoll, Jürgen Wegge, Christine Unterrainer, Silvia Silva, and Thomas Jönsson. Is our knowledge of voice and silence in organizations growing? building bridges and (re)discovering opportunities. *German Journal of Human Resource Management: Zeitschrift für Personalforschung*, 30(3-4):161-194, August 2016. ISSN 2397-0030. doi:[10.1177/2397002216649857](https://doi.org/10.1177/2397002216649857). URL <http://dx.doi.org/10.1177/2397002216649857>.
- [80] Mona Weiss and Elizabeth W. Morrison. Speaking up and moving up: How voice can enhance employees' social status. *Journal of Organizational Behavior*, 40(1):5-19, January 2018. ISSN 1099-1379. doi:[10.1002/job.2262](https://doi.org/10.1002/job.2262). URL <http://dx.doi.org/10.1002/job.2262>.
- [81] Elizabeth Wolfe Morrison. Employee voice and silence: Taking stock a decade later. *Annual Review of Organizational Psychology and Organizational Behavior*, 10(1):79-107, January 2023. ISSN 2327-0616. doi:[10.1146/annurev-orgpsych-120920-054654](https://doi.org/10.1146/annurev-orgpsych-120920-054654). URL <http://dx.doi.org/10.1146/annurev-orgpsych-120920-054654>.
- [82] John Sherry. The complexity paradigm for studying human communication: A summary and integration of two fields. *Review of Communication Research*, 3:22-65, 2015. ISSN 2255-4165. doi:[10.12840/issn.2255-4165.2015.03.01.007](https://doi.org/10.12840/issn.2255-4165.2015.03.01.007). URL <http://dx.doi.org/10.12840/issn.2255-4165.2015.03.01.007>.