

読解時の視線認識情報と音声発話コーパスを用いたポーズの予測とその比較

林小春¹ 狩野芳伸¹¹ 静岡大学 情報学部

{khayashi, kano}@kanolab.net

概要

人間が連続した文章を発話する際、文章中に音声的なポーズが発生する。また、ポーズの挿入は発話時だけではなく音読や黙読といった読解時にも同様に行われ、人間は頭の中で文章にポーズを挿入しながら読解していると考えられる。そこで本研究では、文章読解時の視線認識情報データと音声発話コーパスの二種類を用いて様々な訓練と評価の組み合わせを作り、BERTのファインチューニングにより文章中におけるポーズの適切な挿入箇所を予測した。その結果とデータを比較することで、話し言葉と読解では大きくポーズの挿入パターンが異なることや、各データ内でも個人差やポーズの性質差があり、異なる非均衡パターンを生んでいるとの示唆を得た。

1 はじめに

音声言語で言葉を伝達する際、ポーズ（音声的な間）の挿入により文章が分かりやすくなる箇所がある。ポーズは発話時だけではなく音読や黙読といった読解時にも利用され、文章に明示されていなくても人間は頭の中で文章にポーズを挿入しながら読解していると考えられる。また、人間同士でのコミュニケーションが成立することから、ポーズの適切な挿入箇所は大多数の人間で一致すると考えられる。

発話理解や読解におけるポーズの役割は構文・意味的な区切りのマーカーで、曖昧性解消に寄与すると考える。文章中のポーズ位置予測は音声合成システムにおける特徴量として活用でき、音声の自然さや分かりやすさが向上しうる。テキストベースでも、暗黙のポーズを推測しながら処理することで言語処理に必要な情報を取り込めるため、将来的には言語モデルの学習や推測に資すると期待できる。

多くの文章には読点が含まれるため、読点ごとにポーズを入れるといったルールベースの手法も考え

られるが、それだけでは音声的に適切な位置にポーズが挿入されるとは言えない。

表 1 ポーズを含まない発話文例と含む発話文例

元の発話文	黒い目のきれいな女の子
ポーズを挿入した発話文	黒い / 目のきれいな女の子 黒い目の / きれいな女の子 黒い目のきれいな / 女の子

表 1 の上段に示す発話文例に対して、少なくとも 3 か所にポーズの挿入が可能である。表 1 の下段に示した、ポーズ（/ で表す）を挿入した 3 通りを比較すると、ポーズの挿入位置によって構文構造の解釈が変わり、文章の意味が違ってくる。なおこの例では、同時に複数のポーズを挿入すると曖昧性が解消できず、かえって分かりにくくなる。このような曖昧性解消に資する場合以外にも、ポーズの適切な挿入で音声発話が聞き取りやすくなったり、読解しやすくなることもある。

文章中のポーズ予測に関する研究では、LSTM などの深層学習を用いた手法が、英語を主な対象として提案されてきた [1, 2, 3]。また近年の深層学習手法の発展に伴い、BERT[4] などを用いたポーズ位置の予測が英語 [5] と日本語 [6] を対象に報告されている。この日本語を対象としたポーズ予測では、音声合成の品質向上のためのポーズ予測を目的に、様々な特徴量、モデル構造の効果を検証している。

本研究では、「間」としてのポーズの役割に焦点を置き、適切なポーズ挿入位置とその予測について探求するとともに、音声発話と読解とでポーズの挿入に差異があるのか探求する。そのために、音声発話におけるポーズ情報として日本語話し言葉コーパス (CSJ) [7] を、読解時の情報として現代日本語書き言葉均衡コーパス (BCCWJ) [8] の一部に付与された読解時の視線認識情報データ [9] (以下では BCCWJ-EyeTrack と呼ぶ) を用いる。この二つのデータから訓練・評価の組を組み替えて事前学習済み BERT をファインチューニングし、ポーズ位置

の推測とその結果の比較分析をする。

2 関連研究

2.1 日本語話し言葉コーパス (CSJ)

日本語話し言葉コーパス (Corpus of Spontaneous Japanese, CSJ) [7] は、日本語の自発音声を大量に収集し多くの研究用情報を付加した話し言葉研究用のデータベースであり、質・量ともに世界最高水準の日本語話し言葉データベースである。

その内容は収録した数百時間規模の音声データ (うち約 90% はモノログ音声で、残り約 10% は、対話、朗読、再朗読の音声) を基盤として、すべての音声データを書き起こした転記テキスト、すべての転記テキストを長短 2 種類の語に区切り品詞情報を付与した形態論情報 (短単位と長単位の二種類がある) などの多様な付随データから成っている。

2.2 BCCWJ-EyeTrack

BCCWJ-EyeTrack は、現代日本語書き言葉均衡コーパス (BCCWJ: Balanced Corpus of Contemporary Written Japanese) [8] 中の新聞記事データ 24 件に対し、24 名の被験者により視線走査法を用いてデータを収集したものである [9]。ここでいう視線走査法とは、被験者がディスプレイ画面のどの文字を注視しているかを取得する視線走査装置 (EyeLink1000) を用いて、視線注視箇所と注視時間を計測する手法である。読み戻しなどのより自然な読み時間を取得できるとされており、このデータでは日本語を横書きで表示し、被験者の右目を対象に横方向の動きをミリ秒単位で計測している。データセットは、表 2 の要素から構成されている。

表 2 BCCWJ-EyeTrack の要素

fft(First Fixation Time)	最初に視線が滞留した注視時間
fpt(First-Pass Time)	最初に視線が入って出るまでの時間
spt(Second-Pass Time)	2 回目以降の注視領域に停留する時間
rtp(Regression Path Time)	総注視時間
total(Total Time)	注視領域に視線が滞在する総注視時間

人間が文章を読む際、視線は連続的に移動するのではなく、何文字かを一つの領域として、その領域の間をジャンプする (サッケード) ように非連続的に移動する。人間の網膜は、神経細胞が密集した中

心視野を構成する部分を中心として、周囲に神経細胞がまばらな周辺視野に対応する部分が広がっており、周辺視野の解像度では文字の認識が困難であるため、中心視野で捉えられる範囲のみで文字を認識していると考えられる。中心視野の範囲は視野角で数度程度であり、一度の滞留で認識できる文字の領域は数文字程度ということになる [10]。この非連続的な視線の動きは通常、文章の流れに沿って冒頭から続きへと、短時間の視線滞留とサッケードを繰り返していき、中心視野で捉えた範囲を繋ぎ合わせた文字が認識できることになる。改行をまたいで前の行の末尾から次の行の冒頭に移る場合は、横書きの文書であれば縦方向の移動と横方向の逆行が必要になる。ただし、構文や意味の再解釈が必要になった場合など、何らかの理由で通過済みの場所に視線が戻ることもある (バックトラック)。

2.3 音声合成のための句境界予測

二又ら [6] は日本語音声合成の品質向上のため話し言葉の句境界予測実験を行い、様々な特徴量、モデル構造の効果を検証している。ここでの句境界とは、本研究のポーズとおおむね同等のものと考えられる。自動評価実験の結果、BERT を用いたポーズ予測モデルを導入することにより、その他の予測モデルより大きく予測精度が上昇した。また、それを導入した音声合成に対し主観評価実験を行った結果、音声の自然性も向上することを示した。

実験には CSJ と独自作成の音声コーパスを用い、独自コーパスによる評価結果では BERT と CRF を組み合わせたモデルが F1 値において最も高い数値であったが、BERT のみにおける評価値との差はわずかであり、また CSJ による評価結果では BERT のみのモデルが F1 値において最も高い数値であった。

3 データセット

3.1 音声発話によるポーズデータ

話し言葉のデータセットとして複数話者による対話 (以下、会話と呼ぶ) と単一話者による独話のデータ (以下、独話と呼ぶ) を含む日本語話し言葉コーパス (CSJ) [7] を使用した。CSJ の短単位 (SUW) の形態論情報と、全データ中のコアデータのみが付与されている音声情報から、200 ミリ秒以上の無音区間をポーズとし、その直前のトークンにポーズ有りのラベルを付与した。

表3 ポーズと見なした CSJ のタグと除外した CSJ のタグの一覧

記号	用法	例	処理
<H> <息> <笑>	母音の非語彙的延長 呼吸音 笑い (同時に喋ってはいない)	ソレデ<H> アルワケデ<息> ソウデスネ<笑>	ポーズと見なす
(D), (D2) (F)	断片化した語 フィラー、感情表出系感動詞	(D こ) これ, これ (D2 は) が (F あの), (F うわ)	対象から除外

形態論情報に含まれるタグから、母音の非語彙的延長を表す<H>が付与されているトークンもポーズ有りで見なした。また、<笑>、<息>といった発話と同時に進行していない非言語的表現タグ (表 3) を除外し、200 ミリ秒を超えるものに対してはその直前のトークンをポーズ有りとした。さらに、フィラー・感動詞・断片化した語などのタグ (表 3) が付与されたトークンを除くことで、より内容語のみ連続する文章を抽出できるようにした。

複数話者の会話データに関しては、話者を示すタグを用いて話者ごとの発話に並べ替えた。話し言葉には明示的な文末がないため、直前が助詞でない終助詞の場合や活用形が終止形で直後に助詞や助動詞がない場合などのパターンにより文末を判定し、入力単位の区切りとした。文末には一律ポーズ有りのタグを付与した。先行研究 [6] と同等の設定になるよう、最大長 256 トークン (BERT のサブワード単位) に制限し、最大長を超える文は除外した。抽出したデータセットの統計は表 4 の通りである。

3.2 読解時視線認識によるポーズデータ

書き言葉の読解におけるデータセットとして、BCCWJ の一部に対する視線認識情報データ (BCCWJ-EyeTrack) [9] を使用した。

本研究でのポーズ予測に利用するために、以下の方法でポーズを検出した。まず、データ上同じ領域 (文字のかたまり) に 2 回以上視線が滞留することがあるが、一度目の読解のみを対象とするために、その領域に視線が滞留した合計時間 (total) から二回目以降の滞留時間 (spt) を引いた時間を領域の視線滞留時間とし、各領域に対して計算を行った。

音声発話データにおけるポーズの閾値を 200 ミリ秒に設定したため、各領域の視線滞留時間が 200 ミリ秒を超えるものをポーズと見なし、対応するテキスト末尾のトークンにポーズ有りのラベルを付与した (以下、BCCWJ₂₀₀ と呼ぶ)。しかし、閾値を 200 ミリ秒とした場合、CSJ に比べて目視でもポーズが多すぎる傾向が見られたため、閾値を 300 ミリ秒と

して同様のデータを構成した (以下、BCCWJ₃₀₀ と呼ぶ)。これらは句点でテキストを区切り、BERT への入力単位とした。抽出したデータセットの統計は表 4 のようになった。

表4 各コーパスの入力文 (サンプル) 数

	Train	Val	Test
CSJ	12006	1501	1501
BCCWJ	12525	1565	1566

表5 各コーパスに含まれるポーズ数と総数中のポーズ数の割合。CSJ が短単位、BCCWJ が文節単位

	ポーズ無	ポーズ有	ポーズ有の割合
CSJ	393,682	67,423	14.62%
BCCWJ ₂₀₀	59,095	98,449	62.49%
BCCWJ ₃₀₀	80,878	76,666	48.66%

4 ポーズ予測実験

ポーズ予測を、入力トークン系列に対して各トークンの後にポーズが挿入されるべきか否か出力する、系列ラベリングタスクとして設定した。

4.1 予測に用いた言語モデル

先行研究 [6] の結果から、BERT 単独でもよい性能が得られ、データ数不足やラベル不均衡による評価値の揺れを考えると、他のモデルとの性能差は無視できるレベルと考えられる。本研究では絶対性能の良し悪しより音声発話と読解時視線認識とでポーズの性質や違いの分析に焦点を置くことも踏まえ、BERT を用いて予測を行い、2 つのデータの組み合わせによる予測結果の違いを比較することとした。

実験においては、事前学習済 BERT モデルとして日本語 Wikipedia で事前学習を行った cl-tohoku/bert-base-japanese-whole-word-masking を使用し、以下に述べる訓練データで系列ラベリングとしてファインチューニングを行った。

4.2 実験設定と結果

CSJ と BCCWJ-EyeTrack のそれぞれを訓練: 検証: 評価=8:1:1 に分割し、訓練データは CSJ・BCCWJ₂₀₀・BCCWJ₃₀₀・CSJ+2 種の BCCWJ の 5 通りとし、それ

ぞれのデータで評価した。また、予測結果を出力したものを比較し、その違いを観察した。評価結果を表 6 に示す。

表 6 ポーズ予測実験の評価結果

訓練	評価	適合率	再現率	F1
CSJ	CSJ	53.71	35.42	42.69
	BCCWJ ₂₀₀	73.08	16.52	26.95
BCCWJ ₂₀₀	CSJ	24.28	49.20	32.52
	BCCWJ ₂₀₀	70.36	82.44	75.93
BCCWJ ₃₀₀	CSJ	35.24	20.17	25.65
	BCCWJ ₃₀₀	75.72	60.20	67.07
CSJ+ BCCWJ ₂₀₀	CSJ	51.95	31.39	39.13
	BCCWJ ₂₀₀	69.16	83.39	75.62
CSJ+ BCCWJ ₃₀₀	CSJ	51.12	33.05	40.15
	BCCWJ ₃₀₀	72.74	62.77	67.39

5 考察

5.1 定量的評価結果の考察

CSJ で学習評価を行った先行研究 [6] と比べると、CSJ を用いた評価値が低い。先行研究は CSJ の非コアデータも用いているが、本研究ではポーズ算出のためタイムスタンプの付与されたコアデータのみを用いており、データ量が一桁少なく学習が十分にできなかった可能性がある。また二値分類の問題として捉えると大きく不均衡なデータで、推測結果のバランスや評価尺度の選択によって大きく結果が変わって見えることにも注意する必要がある。

読解時視線認識データにおいて、ポーズと認定する際の閾値を 200 ミリ秒と 300 ミリ秒とで比較した。閾値を大きくするほどポーズと認定されるトークンが減り、ポーズ有無の比率がより不均衡になるため、推測結果においても CSJ のポーズ比率に近づくことで性能が上がって見える可能性がある。しかし、閾値による差分は単にデータ数の増減だけでなく、閾値付近の判断が難しいサンプルが増減することから予測の難易度や性質が異なる可能性がある。実際、表 6 では同じ CSJ での評価に対し、訓練データが BCCWJ₂₀₀ と BCCWJ₃₀₀ とで大きな性能差がみられない。また BCCWJ₃₀₀ 内での訓練・評価よりも、BCCWJ₂₀₀ 内での訓練・評価のほうがスコアが高く、閾値が低いことでほとんど文節区切りに近くなり難易度が下がったと考えられる。

5.2 読解視線認識データの分析

音声発話のポーズ閾値で頻用される 200 ミリ秒に対し、BCCWJ-EyeTrack の各領域における視線滞留

時間は長いものが多く、音声的なポーズが入りうる数と対応させるには閾値を上げる必要があった。一般には音声発話よりも黙読のほうが速いと考えられ [11]、一度の滞留で認識した文字列を処理するほうが同じ文字列を音声で発話するよりも速いならば、人間の「理解」の速度に何らかの上限があるとする、黙読で速く読めた分「理解」が追いつくまで待たなければならないのかもしれない。

5.3 ポーズの時系列と発話者間の違い

発話者によってポーズの入れ方に差があることは、想像に難くない。ポーズを入れうる場所は相当に共通していても、実際にその場所に入れるかどうかは個人差があるものと考えられる [12, 13, 14]。また、ポーズを入れうる場所すべてに入れるとかえって不自然になってしまうことから、全体でのポーズの頻度や、ポーズ間の時間的・文字列的な距離に制約があることが想像される [15, 16]。本研究のような系列ラベリングによりこうしたポーズ間の相関を学習できる可能性があるが、そのために十分な学習データ量であったかは検討の必要がある。

5.4 具体的な事例

付録の表 7、表 8 に CSJ の、表 9、表 10 に BCCWJ-Eyetrack の正解と予測例（閾値に対応し正解₂₀₀ と正解₃₀₀ の二種類正解がある）を示す。

表 7 の正解には、黙読では発生しづらい音声発話特有の非文節境界ポーズがみられ、BCCWJ で訓練した場合と一致していない。

表 8 では、「そこで」の前後などポーズ挿入可能な位置のいずれを選ぶかの違いがみられるが、正解のポーズには話者特有のスタイルが見受けられる。

表 9 と表 10 の正解では、閾値を 200 ミリ秒とした BCCWJ₂₀₀ で学習した場合、ポーズが多めに出力されていることが分かる。CSJ で学習した場合、全体にポーズが少なく、いずれもデータの非均衡度合いが異なること、またその対応の難しさを示唆している。

6 おわりに

話し言葉コーパスと書き言葉読解時の視線認識情報を用いて、文章中のポーズ位置を BERT で予測した。様々な訓練評価のデータ組み合わせと比較の結果、話し言葉と読解時の差異がわかった。データ不均衡と個人差の吸収が将来的な課題である。

謝辞

本研究は JSPS 科研費 JP22H00804, JP21K18115, JP20K20509, JST AIP 加速課題 JPMJCR22U4, および セコム科学技術財団特定領域研究助成の支援を受けたものです。

参考文献

- [1] Viacheslav Klimkov, Adam Nadolski, Alexis Moinet, Bartosz Putrycz, Roberto Barra-Chicote, Thomas Merritt, and Thomas Drugman. Phrase break prediction for long-form reading tts: exploiting text structure information. **INTER-SPEECH 2017**, pp. 1064–1068, 2017.
- [2] Anandaswarup Vadapalli and Suryakanth V. Gangashetty. An investigation of recurrent neural network architectures using word embeddings for phrase break prediction. **INTERSPEECH 2016**, pp. 2308–2312, 2016.
- [3] Paul Taylor and Alan W. Black. Assigning phrase breaks from part-of-speech sequences. **Computer Speech Language**, Vol. 12, No. 2, pp. 99–117, 1998.
- [4] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In **Proceedings of NAACK-HLT**, Vol. 1, p. 2, 2019.
- [5] 楊棟, 郡山知樹, 猿渡洋. 長文音声合成のための BERT 特徴量を用いたポーズ予測の検討. 日本音響学会研究発表会講演論文集, pp. ROMBUNNO.3–Q–21, 2022.
- [6] 二又航介, 朴炳宣, 山本龍一, 橘健太郎. 日本語テキスト音声合成のための句境界予測モデルの検討. 言語処理学会 第 27 回年次大会 発表論文集, pp. 287–292, 2021.
- [7] 前川喜久雄, 籠宮隆之, 小磯花絵, 小椋秀樹, 菊池英明. 日本語話し言葉コーパスの設計 (<特集> 音声研究関連データベースの動向). 音声研究, Vol. 4, No. 2, pp. 51–61, 2000.
- [8] 前川喜久雄. 代表性を有する大規模日本語書き言葉コーパスの構築 (<特集> 日本語コーパス). 人工知能, Vol. 24, No. 5, pp. 616–622, 2009.
- [9] 浅原正幸, 小野創ほか. BCCWJ-eyetrack——『現代日本語書き言葉均衡コーパス』に対する読み時間付与とその分析——. 言語研究, Vol. 156, pp. 67–96, 2019.
- [10] 苧阪直行. 読みの精神物理学: 有効視野の役割を中心に. 哲学研究, Vol. 48, No. 4, pp. 588–612, 1992.
- [11] 高橋麻衣子, 清河幸子. 読解活動における眼球運動の役割. 認知科学, Vol. 20, No. 4, pp. 470–480, 2013.
- [12] 室町俊貴, 狩野芳伸. 対話履歴と言語・音響特徴を用いた相槌生成タイミングの予測. 人工知能学会 2021 年度全国大会論文集, 2021.
- [13] Toshiki Muromachi and Yoshinobu Kano. Estimation of listening response timing by generative model and parameter control of response substantialness using dynamic-prompt-tune. In **INTERSPEECH 2023**, pp. 2638–2642, 2023.
- [14] 室町俊貴, 狩野芳伸. 生成モデルによる傾聴応答タイミングの推定と動的 prompt-tune を用いた応答スタイルのパラメーター制御. 言語処理学会第 29 回年次

大会 (NLP2023) 論文集, 2023.

- [15] 福野将人, 狩野芳伸. 音響的特徴を用いた応答の使分け・挿入を行う傾聴対話システムの試作. 人工知能学会研究会資料 第 84 回言語・音声理解と対話処理研究会 (SLUD), pp. 92–93, 2018.
- [16] 福野将人, 狩野芳伸. 積極傾聴におけるあいづちのタイミングの予測. 言語処理学会第 27 回年次大会 (NLP2021) 論文集, pp. 1605–1610, 2021.

A 付録

A.1 予測結果の例

表7 CSJのテストデータに対する予測結果(1)

正解：米国科学/学会が歴史上の五/大/科学者に/設定ですって

学習データ	予測
CSJ	米国科学学会が歴史上の五大科学者に設定ですって
BCCWJ ₂₀₀	米国科学学会が/歴史上の/五大科学者に/設定ですって
CSJ+BCCWJ ₂₀₀	米国科学学会が/歴史上の五大科学者に設定ですって
BCCWJ ₃₀₀	米国科学学会が/歴史上の/五大科学者に/設定ですって
CSJ+BCCWJ ₃₀₀	米国科学学会が歴史上の五大科学者に設定ですって

表8 CSJのテストデータに対する予測結果(2)

正解：で/他にはメキシコ/の方にも足を延ばして旅行したことが/ありまして/そこで大変/印象深かったのは/初めてピラミッドに登ったことです。

学習データ	予測
CSJ	で他にはメキシコの方にも足を延ばして旅行したことがありまして/そこで/大変印象深かったのは/初めてピラミッドに登ったことです。
BCCWJ ₂₀₀	で他には/メキシコの/方にも/足を延ばして/旅行したことがありまして/そこで大変印象深かったのは/初めてピラミッドに/登った/ことです。
CSJ+BCCWJ ₂₀₀	で他にはメキシコの方にも足を延ばして旅行したことがありまして/そこで/大変印象深かったのは初めてピラミッドに登ったことです。
BCCWJ ₃₀₀	で他には/メキシコの方にも足を延ばして旅行したことがありましてそこで大変印象深かったのは初めてピラミッドに/登ったことです。
CSJ+BCCWJ ₃₀₀	で他にはメキシコの方にも足を延ばして旅行したことがありまして/そこで/大変印象深かったのは/初めてピラミッドに登ったことです。

表9 BCCWJのテストデータに対する予測結果(1)

正解₂₀₀：大学に/進学する時も、「あなたは論理的に考えるのが得意」と、法学部に/行くよう促したのは/池田先生。

正解₃₀₀：大学に/進学する時も、「あなたは論理的に考えるのが得意」と、法学部に行くよう促したのは/池田先生。

学習データ	予測
CSJ	大学に進学する時も、「あなたは論理的に考えるのが得意」と、法学部に行くよう促したのは池田先生。
BCCWJ ₂₀₀	大学に/進学する/時も、「あなたは/論理的に/考えるのが得意」と、法学部に行くよう促したのは池田先生。
CSJ+BCCWJ ₂₀₀	大学に/進学する/時も、「あなたは/論理的に/考えるのが/得意」と、法学部に行くよう/促したのは/池田先生。
BCCWJ ₃₀₀	大学に/進学する時も、「あなたは論理的に考えるのが得意」と、法学部に行くよう促したのは池田先生。
CSJ+BCCWJ ₃₀₀	大学に/進学する時も、「あなたは論理的に考えるのが得意」と、法学部に行くよう促したのは池田先生。

表10 BCCWJのテストデータに対する予測結果(2)

正解₂₀₀：「生意気という/人もいたけれど、/私は、彼女の/ように/モノを/はっきり/言えることが/これからは大切だと/思っていました」。

正解₃₀₀：「生意気という/人もいたけれど、/私は、彼女のようにモノをはっきり/言えることがこれからは大切だと/思っていました」。

学習データ	予測
CSJ	「生意気という人もいたけれど、私は、彼女のようにモノをはっきり言えることがこれからは大切だと思っていました」。
BCCWJ ₂₀₀	「生意気という/人もいたけれど、/私は、彼女のようにモノを/はっきり/言えることがこれからは大切だと/思っていました」。
CSJ+BCCWJ ₂₀₀	「生意気という/人もいたけれど、/私は、/彼女の/ようにモノを/はっきり/言えることがこれからは大切だと/思っていました」。
BCCWJ ₃₀₀	「生意気という/人もいたけれど、私は、彼女のようにモノをはっきり言えることがこれからは大切だと思っていました」。
CSJ+BCCWJ ₃₀₀	「生意気という/人もいたけれど、私は、彼女のようにモノをはっきり言えることがこれからは大切だと思っていました」。