

語りの傾聴において表出する応答データの拡充

村田 匡輝[†] 大野 誠寛[‡] 松原 茂樹^{††}

[†]豊田工業高等専門学校 情報工学科 [‡]東京電機大学 未来科学部

^{††}名古屋大学 情報連携統括本部

murata@toyota-ct.ac.jp

1 はじめに

語ることは人間に備わる基本的な欲求である。語るという行為は、聴き手がいて初めて成立する。社会の個人化が進み、聴き手不在の生活シーンが増加している。人が語れる機会を増やすことは現代社会の重要な課題といえる。

上述の課題に対して、コミュニケーションロボットやスマートスピーカなどの情報機器が語りを聞く役割を担うことが考えられる。機器が聴き手として認められるには、語りを傾聴していることを話し手に伝達する機能を備える必要がある。その明示的な手段は語りに応答することであり、具体的にはジェスチャや発話の表出が有力である。以下では、傾聴的な態度を示す目的で語りに応答する発話を傾聴応答と呼ぶ。

傾聴応答の代表は相槌であり、その生成法が提案されているものの [1, 2]、傾聴を示す応答は相槌以外にも存在する。聴き手としての役割を担うことを目的とした対話システムの開発も行われている [3, 4, 5]。これらのシステムでは、システムが質問や要求を行い、ユーザの話の展開に積極的に関わっていく。一方、本研究が目指す傾聴的応答では、対話の発話権を保持する応答を行うことなく、あくまで一方的な応答を行いつつも、ユーザに傾聴態度を示すことを目指しており、目標とする対話の形態に違いがある。

傾聴を行う情報機器の実現のためには、実際の傾聴応答を収集し、調査・観察を通して傾聴応答の生成方略を検討することが不可欠である。著者らはこれまでに、高齢者の語りに対して作業員 1 名が傾聴応答を付与したデータ（以下、傾聴応答データ）を構築している [6]。本論文では、傾聴応答データの拡充について述べる。作業員を 4 名追加し、計 5 名分の傾聴応答を収集した。傾聴応答は、応答が一つの役割を果たすと考えられる単位に分割し、その種類を分類した。

2 傾聴応答データ

2.1 傾聴応答の収集方法

データの収集は以下の手順で行った。傾聴応答を収集するための対象データとして、高齢者のナラティブコーパス JELiCo¹ を利用した。これは、高齢者の語りの研究を目的に構築されたコーパスであり、30 名の高齢者による平均 20 分の音声データが収録されている。高齢者は用意された 10 個の質問に回答することによって、語りを遂行している。本研究では、上記コーパス 30 名分の高齢者の音声データ（8 時間 42 分 57 秒）を利用して、傾聴応答の収集を実施した。

高齢者の語りの音声に対して、作業員 5 名が聞き役を遂行した。作業員は、接話マイクを使用し、高齢者音声を聞きながら、傾聴を示す応答を行った。本研究では、対象として静的な音声をを用いることで、一つの語りに対して複数作業員による複数パターンの傾聴応答の収集を実現している。

2.2 収集した傾聴応答の整理

収集した応答に対して、書き起こし、形態素情報、形態素単位の発話時間情報を付与した。形態素解析には MeCab (UniDic Ver. 2.1.2) を、発話時間情報の付与には Julius の音素セグメンテーションキットを、それぞれ用いた。また、傾聴応答は、役割を持つと考えられる単位に分割し、人手で役割の種類を分類した。役割は文献 [7] を参考に「相槌」「感心」等、17 種類定めた。

5 名分のデータとして、64,541 個の傾聴応答を収集できた。5 名分の傾聴応答の出現割合を役割ごとに表 1 に示す。相槌が占める割合が最も大きく、67% を越えている。相槌以外の応答は 30% 以上存在し、感心、評価の順に多く出現している。感心は、語りの内容に

¹<https://www.gsk.or.jp/catalog/gsk2018-a/>

表 1: 傾聴応答の出現割合

ラベル	頻度	割合 (%)
相槌	43,508	67.41
感心	11,284	17.48
評価	2,001	3.10
繰り返し応答	1,622	2.51
同意	1,586	2.46
その他	992	1.54
納得	883	1.37
驚き	762	1.18
言い換え	597	0.92
意見	412	0.64
考えている最中	260	0.40
不同意	242	0.37
補完	230	0.36
あいさつ	107	0.17
想起	31	0.05
驚きといぶかり	17	0.03
考え始め	7	0.01

対して、感心や気付き等の態度を示す応答であり、以下のようなものがある。

【語り】孫が六人おりました

【応答】そうですかー

評価は、語りの内容が示す事象に対する態度を示す応答で、

【語り】祖父母に対してですね気を使って

【応答】すばらしいですねー

といったものが挙げられる。また、繰り返し応答は、語りの発話と同じ言葉を用いて行う応答であり、以下の例に示すようなものである。

【語り】蔵王スキーに行ったときですね

【応答】蔵王に

傾聴応答データの例を図1に示す。図1では、語りについては、形態素単位の表層形を発話時間とともに示している。応答については、役割が付与された単位ごとに、作業者5名分(A~E)のものを、発話時間とともに示している。

「歩くのが得意」という高齢者の語りに対して、「そうですか(応答A)」「そうなんですか(応答B)」と

いう応答で感心が示されていたり、「歩くことはいいことですよ(応答D)」「私もです(応答E)」という応答で同意が示されている。また、「一万歩以上歩いている」という語りに対しては、作業者A, B, Dの3人によって「すごいです」という傾聴応答が行われている。しかし、その役割としては、応答の発話のタイミングや発話方法によって、それぞれ評価、驚き、感心と別のものになっている。

5名それぞれの傾聴応答の出現頻度、種類数を役割ごとに表2に示す。種類数とは傾聴応答の文字列の異なりを意味する。同一の語りの音声に対して、作業者Aは16,884回、作業者Eは9,800回の応答を行っている。発話される傾聴応答の頻度、表現形式には個人差があることが分かる。

3 傾聴応答データの調査

収集した傾聴応答データを用いて、作業者の増加に対する傾聴応答の種類数の増加率、および、応答タイミング一致度について調査を行った。

3.1 収集した応答の増加率

本データは、静的な語りに対して複数の作業者が応答を付与することによって構築している。語りに傾聴を示すための応答として、その表現形式や応答表出のタイミングには適切なものが複数存在すると考えられるため、1名分のデータでは十分に応答を収集できていない可能性がある。一方、傾聴応答の役割等はある程度体系立てられており、パターンは無尽蔵に存在するわけではない。本節では、作業者Aの応答を基準に、作業者B以降の応答を追加したときの、役割ラベルごとの種類数の増加率について調査した。

相槌についての結果を図2に示す。作業者Bの追加時点で種類数が1.28倍に増えるが、その後は増加率は減少していき、作業者Eの追加時は1.03倍しか増えていない。収集した相槌の種類数としては収束しつつあると考えられる。一方で、繰り返し応答や言い換え、意見など、応答の対象となる語りの内容に応じて表現形式が異なる応答の種類数は、まだ増加する可能性があると考えられる。

3.2 応答タイミングの一致率

本節では、傾聴応答の発話開始タイミングの一致率について調査する。本データでは、高齢者の語りの音

語り		応答A			応答B			応答C			応答D			応答E		
発話時間	形態素	発話時間	応答	役割	発話時間	応答	役割	発話時間	応答	役割	発話時間	応答	役割	発話時間	応答	役割
1.96 - 2.36	歩く															
2.36 - 2.53	の															
2.53 - 2.64	が															
2.64 - 3.03	得意															
3.03 - 3.16	な															
3.16 - 3.22	ん															
3.22 - 3.44	です	3.43 - 4.34	そうですか	感心												
3.44 - 3.49	sp															
3.49 - 3.67	ね				3.59 - 3.73	あー	感心									
3.67 - 3.87	sp				3.76 - 4.63	そうなんですか	感心				3.84 - 3.99	あ	感心			
3.87 - 4.14	pause										4.07 - 5.55	歩くことはいい ことですよね	感心 同意	3.91 - 3.95	あ	感心
4.14 - 4.98	毎日							4.60 - 4.79	はい	相槌				3.95 - 4.71	私ものです	同意
4.98 - 5.20	ー															
5.20 - 5.51	万	5.37 - 5.57	えー	相槌												
5.51 - 5.67	歩															
5.67 - 6.16	以上															
6.16 - 6.35	sp															
6.35 - 6.68	歩い	6.46 - 7.78	すごいですね	評価	6.41 - 6.79	はー	感心									
6.68 - 6.76	て															
6.76 - 6.89	おり															
6.89 - 7.25	ます				7.19 - 8.33	すごいですね	驚き				7.17 - 7.32	や	驚き	7.23 - 7.28	あ	感心
7.25 - 7.45	sp										7.37 - 8.49	すごいですー	感心			
7.45 - 7.73	pause							7.31 - 7.55	はい	相槌				7.57 - 8.14	いいですね	同意

図 1: 傾聴応答データの例

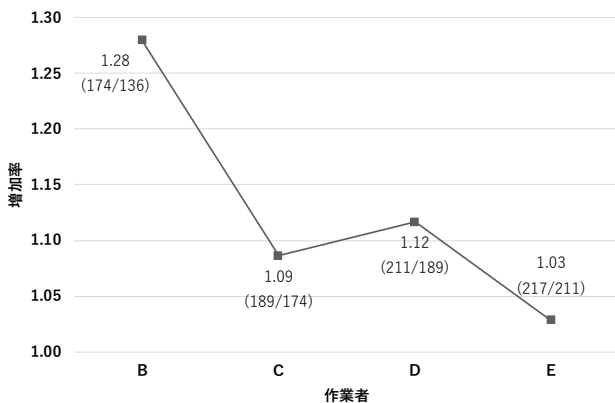


図 2: 相槌の種類の増加率

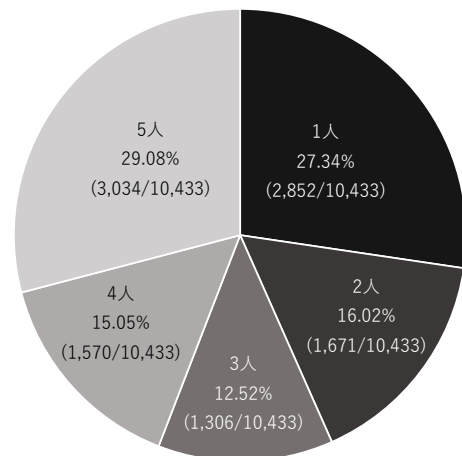


図 3: 応答タイミングの一致率

声を聞きリアルタイムで作業者が応答を行っているため、同じ語りに対して行われた応答であっても、発話開始時間にずれが生じる。本研究では、文節境界を応答開始可能なタイミングの候補とし、実際に行われた応答を、語り、応答の発話時間を利用して対応付けた。具体的には、応答の第一形態素の発話開始までに発話が終了している語りの形態素を末尾に持つ文節の直後へと対応付けた。

調査には、収集した応答の一部である 38,521 個の応答を用いた。対応付けの結果、応答は 10,433 個の文節境界に対して対応付けられた。調査に用いたデータ

中には 17,476 個の文節境界が存在したため、59.70% の文節境界で応答が行われたことになる。

応答の開始タイミングが一致した人数とその割合を図 3 に示す。応答が行われた文節境界のうち、72% 以上は、複数作業者が応答している。応答の開始タイミングはある程度一致することが分かる。また、1 人のみが応答している文節境界 2,852 個のうち、各作業者の割合はそれぞれ、A が 43.30%、D が 19.42%、C

表 2: 作業者毎の傾聴応答の出現頻度と種類数

ラベル	A		B		C		D		E	
	出現頻度	種類数	出現頻度	種類数	出現頻度	種類数	出現頻度	種類数	出現頻度	種類数
相槌	10,169	136	8,227	106	10,466	87	8,440	99	6,206	53
感心	3,077	241	2,296	157	1,228	102	2,626	272	2,057	143
評価	680	358	160	59	210	68	440	217	511	115
繰り返し応答	753	672	0	0	91	73	664	581	114	104
同意	572	127	164	40	220	50	313	101	317	63
その他	137	57	64	39	193	94	556	218	42	22
納得	367	99	101	24	62	11	90	29	263	37
驚き	176	26	134	38	126	37	194	55	132	24
言い換え	156	152	0	0	20	17	358	331	63	56
意見	265	239	3	3	34	29	89	82	21	20
考えている最中	251	23	0	0	0	0	9	4	0	0
不同意	105	48	32	13	15	4	68	23	22	12
補完	110	109	0	0	2	0	88	83	30	26
あいさつ	28	10	29	9	13	9	23	12	14	6
想起	18	16	0	0	2	0	3	3	8	4
驚きといぶかり	16	5	1	1	0	0	0	0	0	0
考え始め	4	3	3	2	0	0	0	0	0	0
合計	16,884	2,321	11,214	491	12,682	581	13,961	2,110	9,800	685

が 15.88%, B が 14.31%, E が 7.08% である。応答の頻度が多くなれば, 1 人のみが行う応答の割合も当然ながら増加する。

応答が同じタイミングで行われていても, 発話されている応答の役割は異なる。複数の作業者が応答を行っているタイミングにおけるラベルの一致具合の調査は今後の課題である。

4 おわりに

本論文では傾聴性を備えた情報機器の実現のためのデータとして, 傾聴応答データの収集について述べた。作業者 1 名分の既存データに加え, 新たに作業者 4 名分の応答を収集し, 計 5 名分の傾聴応答データへと拡充した。合計で 64,541 個 (4,981 種類) の応答を収集した。

今後, 引き続き別の作業者による応答を収集し, さらなるデータの拡充を目指す。また, 拡充したデータを用いて, 応答タイミングの推定や傾聴応答の生成に取り組む。

謝辞 高齢者のナラティブコーパスは, 奈良先端科学技術大学院大学ソーシャル・コンピューティング研究室から提供いただいた。本研究は, 一部, 科学研究費補助金 (挑戦的研究 (萌芽)) (No. 18K19811) により実施したものである。

参考文献

- [1] 大野誠寛, 神谷優貴, 松原茂樹. 対話コーパスを用いた相づち生成タイミングの検出. 電子情報通信学会論文誌, Vol. J100-A, No. 1, pp. 53–65, 2017.
- [2] 山口貴史, 井上昂治, 吉野幸一郎, 高梨克也, Nigel G. Ward, 河原達也. 傾聴対話システムのための言語情報と韻律情報に基づく多様な形態の相槌の生成. 人工知能学会論文誌, Vol. 31, No. 4, pp. C–G31.1–10, 2016.
- [3] 下岡和也, 徳久良子, 吉村貴克, 星野博之, 渡部生聖. 音声対話ロボットのための傾聴システムの開発. 情報処理学会論文誌, Vol. 24, No. 1, pp. 3–47, 2017.
- [4] 目黒豊美, 東中竜一郎, 堂坂浩二, 南泰浩. 聞き役対話の分析および分析に基づいた対話制御部の構築. 情報処理学会論文誌, Vol. 53, No. 12, pp. 2787–2801, 2012.
- [5] 小林優佳, 山本大介, 土井美和子. 高齢者対話インタフェース —発話間の共起性を利用した傾聴対話の基礎検討—. FIT2011 講演論文集, Vol. 10, No. 2, pp. 253–256, 2011.
- [6] 村田匡輝, 大野誠寛, 松原茂樹. 語りの傾聴を話し手に示す応答発話の収集. 電気学会論文誌 C, Vol. 138, No. 5, pp. 637–638, 2018.
- [7] 日本語記述文法研究会. 現代日本語文法 7. くろしお出版, 2009.