

日本語の発話における談話標識の一考察 —テキストマイニング手法と目視による分析を通して—

A Study on Discourse Markers in Oral Japanese Analyzed through Text-Mining and Observation

原田 朋子

要 旨

本研究の主たる目的は、話し言葉の談話における談話標識のうち、書き言葉で見られるような接続表現や副詞の意味・用法が話し言葉の中で弱まり、いわばフィルター化している実態を明らかにすることであった。そのために、会話のコーパスを対象に、テキストマイニング手法により、まず発話者のフィルターの使用傾向について、頻度集計したのち、主に発話者の年代という位相の観点から、特徴抽出を行った。その上で、接続表現と副詞に関しても同様の手法で特徴情報の抽出を行い、ある年代で頻繁に使用される傾向のあるフィルター、接続表現、副詞を明らかにした。テキストマイニングによる統計的分析では、それらの使用状況に関して、視覚化することができ、発話者の年代別使用傾向の特徴をより客観的に示すことができた。

さらに、目視による質的分析で、ある一定の接続表現や副詞が談話の中でフィルター化している現象を明らかにした。例えば、「でも」「で」「だって」のような接続表現について、本来の意味・用法が失われ、フィルター化しているものが一部存在する一方で、「だから」のように、話し言葉の中で多用されていても、フィルター化しにくい接続表現もあることが判明した。そして、副詞については、「まあ」や「もう」などはフィルター化しやすく、「やっぱ」「やっぱり」「結構」などはフィルター化しやすいとまでは言えないものの、フィルター化するものを見出した。

キーワード

日本語 発話 談話標識 テキストマイニング フィルター 接続表現 副詞

1 はじめに

書き手や話し手の談話展開の手がかりとなる談話標識には、さまざまな要素があるが、中でも、接続表現や副詞は、書き言葉にも話し言葉にも使用され、接続表現は、とりわけテキスト形成に関わる機能を有し、接続表現によって文脈展開が創出される。また、副詞には、主に用言や述部、文に対する修飾限定のはたらきがある。一方で、フィ

ラーは、話し言葉で多用され、話し手や聞き手の心的態度を表す傾向がある。しかしながら、接続表現や副詞に、その本来の意味や用法が時に失われているようなものが、話し言葉の中に見受けられることがある。

そこで、本稿では、会話のコーパスを対象に、まず、そのコーパスにおける発話者のフィルターの使用状況を頻度集計したのち、発話者の年代別使用傾向の特徴を見出す。

次に、同コーパスにおいて、接続表現と副詞を抽出し、それらについても頻度集計を実施した上で、発話者の年代別の特徴を明らかにする。年代別という属性について、ことさら特徴抽出にこだわるのは、一般的に、若者の発話において、接続表現や副詞の元来の意味・用法が弱まった使用傾向があると言われることがあるが、果たしてそうだろうかという疑問を抱いたことにある。

さらに、本研究では、接続表現と副詞に関して、主に発話頭に出現するものに着目し、それらが元来の意味や用法を保持して使用されているかについて、目視による考察を試みる。

筆者は、原田（2019）において、書き言葉を対象に、日本語母語話者と日本語学習者の接続表現の使用実態をテキストマイニング手法と目視による分析を通して実施した。頻度集計、特徴抽出等について、計量的観点から統計的に分析する際に、テキストマイニング手法を用いると、より大量のデータ分析が可能となり、分析結果の再現性の確保の面からも、客観的な結果が得られ、且つ、得られたデータを俯瞰することができる。また、テキストマイニングの多変量解析により、多次元のことを視覚化することも可能となる。ただ、孤立した文の中に留まらず、文や発話を有意義に結びつけていくような、談話の中での言語形式を考察する質的研究においては、目視による分析も不可欠だと考えるため、テキストマイニングと目視の両面からの分析を行おうとするのが筆者の立場であり、本稿の分析方法の特徴である。

2 フィラーに関する先行研究

筆者はこれまで、日本語の主に書き言葉を対象とした研究を行ってきた。しかし、書き言葉と話し言葉では、表現の形態的な差もさることながら、特に話し言葉には、意味や用法が曖昧な用いられ方の接続表現が多く、また、発話頭で用いられやすい一定の副詞が見られる。そのことについては、後の章で述べるが、本研究では、まず、話し言葉特有の談話標識の一つであるフィルターに着目する。

話し言葉を対象にしたフィルターに関する先行研究は数多く存在する。一般的に挙げられる「あの」「あのー」「まー」「えー」などに関して、「場つなぎ言葉」、「間つなぎ言葉」、「言いよどみ」、「無意味語」のようなさまざまな言い方がなされるが、本稿では、定延（2016）や多数の先行研究のように「フィルター」と呼ぶ。

また、中島（2011）は、フィルターについて、本来の語彙的な意味から離れて用いられ、それを取り去っても文・談話全体の命題的な意味や内容が変わらないような表現と定

義している。

そして、中島 (2011) は、1 対 2 以上の対話の話し言葉、それも自然談話に現れるフィラーを対象とした調査で、各々のフィラーの場面差および世代差について、数値を表示し、以下のように述べている。

自然談話 6,000 発話から採取されたフィラーは、総数 1,630 件あり、最も出現率が高いフィラーは、コソア型の「アノー」「アノ」で 32.7% (対総数比、以下同様) を占める。以下「ア」(8.3%)、「デ」(7.7%)、「ナンカ」(7.1%)、「エー」(5.3%)、「マー」(5.3%)、「エート」(4.6%) という順位となっている。

さらに、中島 (2011) は、自然談話のフィラーは、その出現位置が発話頭か発話中か発話末かによって、談話の中で果たす機能が異なり、発話頭に出現するフィラーは、発話境界の明示、発話の切り出し、発話権の維持、前の発話の補正、話し手の心的態度の表出等の機能を持つとし、発話中に出現するフィラーは、発話展開に関与する機能として間つなぎ語、注意喚起、話し手の心的態度の表出等といった機能を持つとしている。そして、発話末に出現するフィラーは、言いよどみ・言いさし、発話終了という機能を持つと述べている。但し、フィラーの中には上記の機能を併せ持ったり、重なり合ったりするものがあり、個々の解釈にはゆれが生ずることがあるという。

他にも、山根 (2002) は、テレビのトーク番組の対話とインタビュー番組の対話を対象に、フィラーの種類、音声面、発話・談話上の出現位置、役割、属性との関わりなど、様々な角度からフィラーの分析を行っている。そこには、進行役とゲストという話し手の関係性があり、いかに相手との関係を壊さず、話を展開させていくかに注意が払われると述べている点が、雑談のような自然談話を対象とした場合と特徴が異なると思われる。

山根 (2002) の 8 つの資料におけるフィラーの種類は、全 49 種類、全出現数は 1,394 例である。出現数が多いのは、「アノ」314 例 (22.5%)、「アノー」205 例 (14.7%)、「マ」124 例 (8.9%)、「マー」112 例 (8.0%) で、これらで全出現数の 5 割強を占めるという。型別の傾向では、コソア型の頻度が傑出している点の中島 (2011) の結果と共通している。その他のフィラーについては、モー型が 103 例 (7.4%)、コーソー型が 96 例 (6.9%)、ナンカ型が 84 例 (6.0%) あり、特にナンカ型に関しては種類の多さが目立ったという。なお、ゲストに関西方言話者がいたにもかかわらず、方言形のフィラーは全く出現していないことについて、テレビというメディアを通しての対話であり、テレビの前の聴衆を意識しているためではないかと言及している。出現位置からの考察では、発話の途中の出現位置として、助詞の後、フィラーの後、副詞の後の順に出現頻度が高く、接続詞の後には「マー」が最も多く出現し、接続詞との結びつきが強いことがうかがえると指摘している。また、フィラーは、その多くが意味の切れ目の大きい位置に出

現しやすく、語の結びつきの強い位置には出現しにくいことが分かるとしている。

その他にも、山根（2002）は、フィラーと属性との関わりについて、性差、進行役とゲストという立場、年齢差、個人差といった点から考察しているが、発話時間との相関関係から、1回フィラーが現れるまでの時間が男性4.8秒、女性5.6秒と、男性の方がフィラーの使用頻度が高いと述べ、年齢の低いゲストとの対話に「アノー」「コ（ー）」「ソノ（ー）」「ナンカ」「マー」の頻度が高いと指摘している。

3 テキストマイニングと目視による分析

3.1 分析対象とテキストマイニングによる分析方法

本研究の分析対象は、『名大会話コーパス』¹で、129 会話、合計約 100 時間の日本語母語話者同士の雑談を文字化したコーパスである。会話の参加者の出身地は、29 都府県（愛知県 32 人、東京都 21 人、岐阜県 14 人、神奈川県 8 人、千葉県・静岡県 各 7 人、大阪府・兵庫県 各 5 人、山梨県・岡山県 各 4 人、宮城県・滋賀県・福岡県 各 3 人、三重県・京都府・山口県・熊本県 各 2 人、福島県・山形県・茨城県・埼玉県・群馬県・福井県・島根県・愛媛県・大分県・宮崎県・鹿児島県・沖縄県 各 1 人）という内訳となっている。また、参加者は、男性 22 人（20 代前半～60 代後半）、女性 114 人（10 代後半～90 代前半）である。図 1 に、男女別・年代別の延べ参加回数を視覚化した。

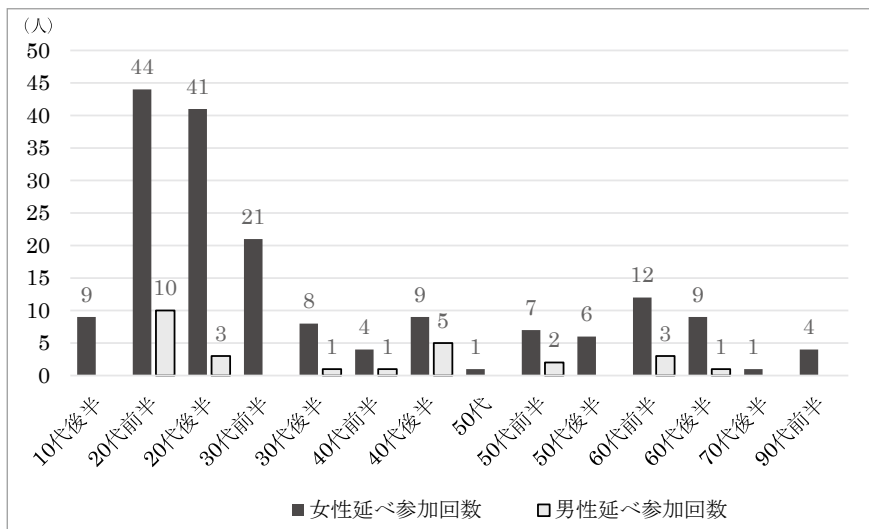


図 1 男女別・年代別の延べ参加回数

図 1 から明らかなように、参加者は 20 代前半から 30 代前半の女性が圧倒的に多いと言える。また、男性の中では、20 代前半の参加者が最も多く、次に 40 代後半の参加者が比較的多く見られる。

『名大会話コーパス』のデータは、可能な限り音声に忠実に文字化されたテキストファイル形式で提供されている。129 会話収録されているが、本研究対象は、このうちの

100 会話である。『名大会話コーパス』はテキスト化される際に、聞き取れない部分や相づちの部分等について、記号や括弧等に表示されている。そのため、テキストマイニングによる分析を実行するにあたり、それらが障害となることから、テキスト中の記号や括弧等については、以下の方針により削除（データクリーニング）した。

- (1) 聞き取れない部分に付けられた***表示を削除
- (2) 相づちと判断され、括弧が付与されている箇所の相づちを残し、() を削除
- (3) 会話中の笑いを表示する<笑い>を削除
- (4) 各発話の初めに付けられた発話者コードについては、発話部分から削除
- (5) 一定の長さ以上の沈黙を表すために付けられた<間>表示を削除
- (6) 漢字の発音が問題になりそうなものに付けられた【かな】表示を削除
- (7) 発話部分ではない補足情報を削除

上述の手順の後、金明哲氏によって開発された MTMineR (Multilingual Text Miner with R) を用いて、テキストマイニングによる複数の分析を行った。MTMineR は、テキスト型データを構造化して集計し、 R^2 を用いて統計的に分析するソフトウェアである。

3.2 フィラーの分析

3.2.1 頻度集計

全てのテキストの中からフィラーを抽出するために、まず MeCab により形態素解析を実施し、そのうちフィラーを MTMineR の n-gram を用いて抽出し、集計したものが表 1 である。

表 1 フィラーの出現回数

	出現回数		出現回数		出現回数
ナンカ	2,830	ソノー	151	エト	17
アノー	1,309	エート	146	ト	17
エ	728	マ	133	マア	4
アー	432	エート	91	ウン	2
アノ	411	エーット	80	ソウデスネ	2
ア	284	ウント	19		

集計の結果、フィラーとして抽出されたものは、17 種類、6,656 例あった。出現回数の多かったものに、「ナンカ」「アノー」「エ」「アー」「アノ」等が現れた。これらのフィラーについて、発話者の属性の違いによる使用状況の特徴や差異があるのかどうかを見るために、以下に出身地別、年齢別による使用状況について分析を試みた。

3.2.2 出身地別のフィルターの使用状況

次に、出身地別のフィルター使用率（発話数／フィルター出現回数）を図2に示した。使用率はいくつの発話に1回フィルターを使用しているかを表した指標で、数字が小さいほど頻繁にフィルターを使用していることになる。図2から、出身地別にフィルターの使用が少ないのは、茨城県 56.9 発話、島根県 53.2 発話、沖縄県 39.5 発話と続いている。一方で、フィルターの使用が多いのは、福島県 2.8 発話、福岡県 6.3 発話、大分県 8.6 発話と続いている。出身地別のフィルター使用率の最大値と最小値の差は約20倍であるが、出身都府県ごとの参加人数が一様でないことから、図2に関しては、ここでのさらなる言及は避けるが、出身地などの属性に関係なく、フィルター使用率の平均は12.4 発話であり、図2のフィルター使用率の分布状況からも、概ね10～20 発話に1回、フィルターを使用していることが分かる。

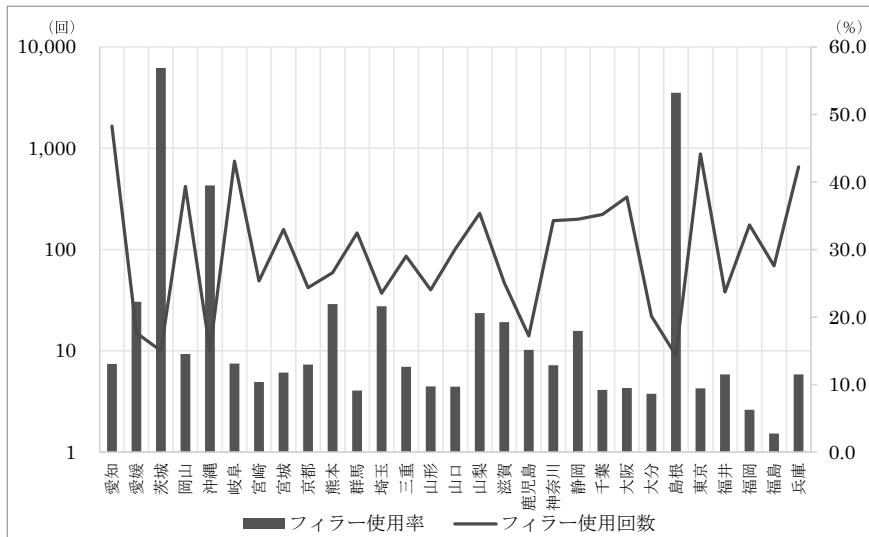


図2 出身地別のフィルター使用率

3.2.3 発話者の年代別のフィルターの使用状況

次に、発話者の年代別の使用状況について分析を行った。『名大会話コーパス』では、15歳～19歳を10代後半として、同様に、20歳～24歳を20代前半、25歳～29歳を20代後半というように、90歳～94歳の90代前半までの5歳区切りで年齢による分類がなされている。但し、例外として、1例だけ50代（50歳～59歳）という分類があるが、この1例については、分析に含めたものの、この1例のみを取り挙げての細かな考察は行わない。

図3は、発話者の年代別のフィルター使用率（発話数／フィルター出現回数）を示したものである。出身地別のフィルター使用状況と同様に、使用率はいくつの発話に1回フィルターを使用しているかを表す指標で、数字が小さいほど頻繁にフィルターを使用してい

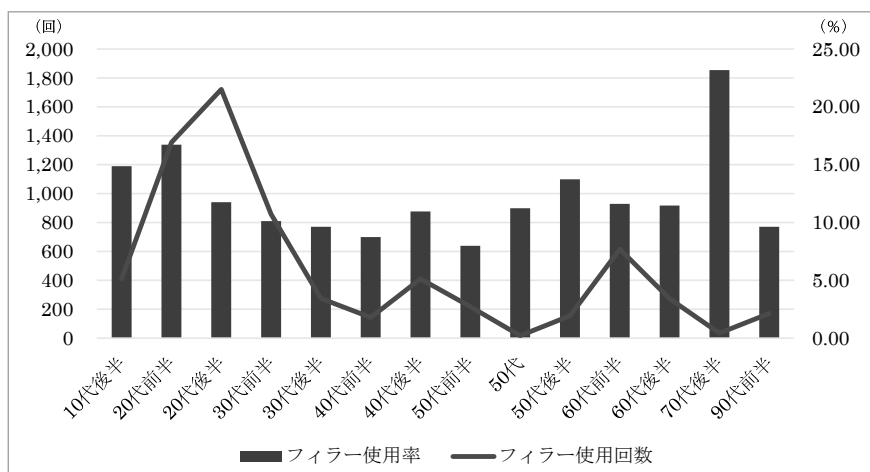


図3 発話者の年代別のフィラー使用率

ることになる。図3から、フィラーの使用が比較的少ないのは、70代後半23.2発話、20代前半16.7発話、10代後半14.9発話と続いている。一方で、使用が相対的に多い年代は、50代前半8.0発話、40代前半8.7発話、30代後半と90代前半9.6発話の順となっている。図3に関する年代別の使用状況の差異についても、コーパスの全体分析対象者の年代別の人数が一樣でないことから、出身地別分析と同様に、図3に関しては、これ以上の言及はしない。

しかし、抽出したフィラーごとの年代別の使用回数については集計を続行した。年代別の各フィラー使用回数を表2に示し、表3には年代別の一人当たりの各フィラー使用回数を示した。

使用回数の最も多い「ナンカ」についてであるが、各年代ともに使用しており、最も多く使用しているのは20代後半の889回で、最も使用が少ないのは70代後半で1回となっている。次に多く見られた「アノー」も各年代で使用されているが、中でも60代前半は264回、20代後半は262回と多く見られる。また、「エ」は20代前半に234回、20代後半に192回であり、20代に多く見られる。他方で、使用回数が最も少なかった「ソウデスネ」は、20代後半と40代後半に1回ずつ出現しているのみであった。

表2 発話者の年代別の各フィラー使用回数

	なんか	あの一	え	あー	あの	あ	その一	えーと	ま	えー	えーっと	うんと	えと	と	まあ	うん	そうですネ
10代後半	238	9	48	34	23	18	2	3	12	14	5	3	0	2	0	0	0
20代前半	486	135	234	160	81	90	35	30	48	32	14	4	2	1	2	0	0
20代後半	889	262	192	113	88	47	38	13	27	27	12	1	4	6	0	1	1
30代前半	535	89	69	31	22	26	13	28	19	4	16	2	2	1	0	1	0
30代後半	158	63	10	10	19	4	3	4	2	0	0	2	0	0	0	0	0
40代前半	117	11	6	2	1	4	0	2	0	0	0	0	0	0	0	0	0
40代後半	126	129	35	9	36	18	18	8	4	1	17	1	6	1	0	0	1
50代前半	51	111	18	3	13	0	8	10	0	0	3	1	0	0	0	0	0
50代	8	6	0	0	0	0	2	0	0	0	1	0	0	0	0	0	0
50代後半	38	67	6	5	9	7	9	5	2	1	3	0	0	3	0	0	0
60代前半	132	264	55	38	44	39	8	20	5	2	5	0	1	2	1	0	0
60代後半	40	93	19	24	37	17	15	15	2	5	3	1	2	0	0	0	0
70代後半	1	4	7	2	4	7	0	0	1	4	0	4	0	0	0	0	0
90代前半	11	66	29	1	34	7	0	8	11	1	1	0	0	1	1	0	0

表 3 発話者の年代別一人当たりの各フィルター使用回数

<出現率>	なんか	あの一	え	あー	あの	あ	その一	えーと	ま	えー	えーっと	うんと	えと	と	まあ	うん	そうですね
10代後半	26.44	1.00	5.33	3.78	2.56	2.00	0.22	0.33	1.33	1.56	0.56	0.33	0.00	0.22	0.00	0.00	0.00
20代前半	9.00	2.50	4.33	2.96	1.50	1.67	0.65	0.56	0.89	0.59	0.26	0.07	0.04	0.02	0.04	0.00	0.00
20代後半	20.20	5.95	4.36	2.57	2.00	1.07	0.86	0.30	0.61	0.61	0.27	0.02	0.09	0.14	0.00	0.02	0.02
30代前半	25.48	4.24	3.29	1.48	1.05	1.24	0.62	1.33	0.90	0.19	0.76	0.10	0.10	0.05	0.00	0.05	0.00
30代後半	17.56	7.00	1.11	1.11	2.11	0.44	0.33	0.44	0.22	0.00	0.00	0.22	0.00	0.00	0.00	0.00	0.00
40代前半	23.40	2.20	1.20	0.40	0.20	0.80	0.00	0.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
40代後半	9.00	9.21	2.50	0.64	2.57	1.29	1.29	0.57	0.29	0.07	1.21	0.07	0.43	0.07	0.00	0.00	0.07
50代前半	51.00	111.00	18.00	3.00	13.00	0.00	8.00	10.00	0.00	0.00	3.00	1.00	0.00	0.00	0.00	0.00	0.00
50代	0.89	0.67	0.00	0.00	0.00	0.00	0.22	0.00	0.00	0.00	0.11	0.00	0.00	0.00	0.00	0.00	0.00
50代後半	6.33	11.17	1.00	0.83	1.50	1.17	1.50	0.83	0.33	0.17	0.50	0.00	0.00	0.50	0.00	0.00	0.00
60代前半	8.80	17.60	3.67	2.53	2.93	2.60	0.53	1.33	0.33	0.13	0.33	0.00	0.07	0.13	0.07	0.00	0.00
60代後半	4.00	9.30	1.90	2.40	3.70	1.70	1.50	1.50	0.20	0.50	0.30	0.10	0.20	0.00	0.00	0.00	0.00
70代後半	1.00	4.00	7.00	2.00	4.00	7.00	0.00	0.00	1.00	4.00	0.00	4.00	0.00	0.00	0.00	0.00	0.00
90代前半	2.75	16.50	7.25	0.25	8.50	1.75	0.00	2.00	2.75	0.25	0.25	0.00	0.00	0.25	0.25	0.00	0.00

表 3 における年代別の一人当たりの各フィルター使用回数は、表 2 のフィルター使用回数を各年代別の延べ参加人数で除して算出しており、1 会話³における 1 人のフィルター使用回数を示したものである。

表 3 では、使用回数 10 以上のものを黒地白抜き表示しているが、「ナンカ」の使用は、10 代後半から 50 代前半までの年代で多く使用されており、50 代後半以上では少ないことが明らかになった。

一方で、「アノー」の使用は、40 代後半以上の年代に多く、10 代後半から 40 代前半までの年代では使用が少ないことが分かる。また、40 代後半では、「ナンカ」「アノー」の使用率はほぼ同率となっている。また、50 代前半では、他のフィルターも多く使用されている。顕著に使用実態に差が表れた「ナンカ」と「アノー」については、年代別使用実態をさらに視覚化し、図 4 と図 5 に漏斗型グラフにより示しておく。

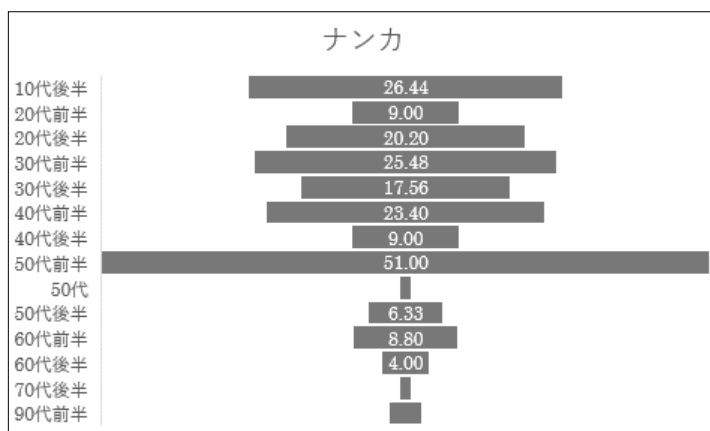


図 4 「ナンカ」に関する発話者の年代別の出現回数

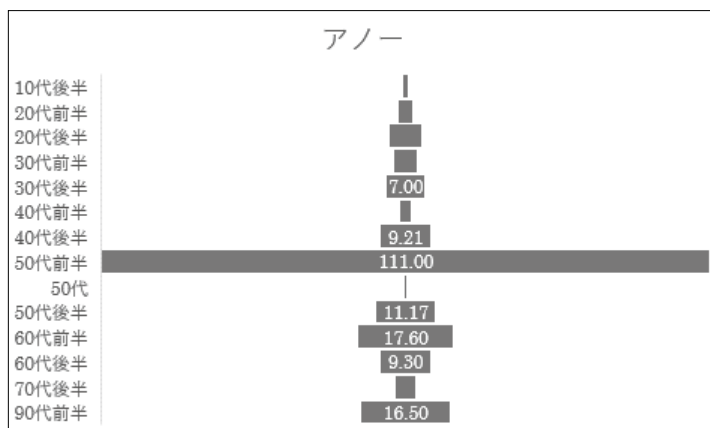


図 5 「アノー」に関する発話者の年代別の使用回数

3.2.4 フィラーの使用頻度状況に関する考察

頻度集計の結果、フィラーとして抽出されたものは、17種類、6,656例あり、出現回数の多かったものに、「ナンカ」「アノー」「エ」「アー」「アノ」等が現れた。

出身地別のフィラー使用状況は、参加人数では愛知県及び東京都出身者の数が多かったにも関わらず、フィラーの使用率は、福島県、福岡県、大分県出身者に比較的多く、茨城県、島根県、沖縄県出身者の使用が比較的少ないことが分かった。

また、年代別のフィラー使用状況では、一人当たりのフィラー使用回数により、フィラーごとの使用状況を分析した。その結果、ある特定の年代層に顕著に使用されるフィラーがあることが明らかになった。具体的には、「ナンカ」は10代後半から50代前半までの年代で多く使用されており、50代後半以上では少ないことや、「アノー」は40代後半以上の年代に多く、10代後半から40代前半までの年代では使用が少ないことが判明した。

なお、形態素解析の結果、フィラーとして抽出されたものの数は多いが、種類がやや少ないように思われるため、今後、フィラーを中心とした分析を行う場合には、形態素解析ツールの辞書に予め先行研究に挙げられているようなフィラーを登録しておくが必要になるだろう。

3.3 接続表現の分析

3.3.1 テキストマイニングによる接続表現の分析

前項では、フィラーに関して、主に頻度の観点からの分析結果を記したが、発話の冒頭や途中や、時には発話末にも現れる談話標識は、言うまでもないことだが、フィラーだけではない。

筆者はこれまでに書き言葉を対象にした接続表現について、頻度や分布や文脈展開に関わる機能等についての研究を行ってきたが、本稿では話し言葉における接続表現の頻度や機能について分析を行う。また、原田(2005)においては、書き言葉(小論文やレポート)の執筆者が予備校講師などの日本語母語話者と中国語母語話者である日本語学習者であり、両者の使用状況の比較を行ったが、本稿においても、年代別の属性の違いによる特徴を見ていく。

フィラーの分析と同様に、全てのテキストの中から接続表現を抽出するために、まずMeCabにより形態素解析を実施し、そのうち「接続詞」をMTMineRのn-gramを用いて抽出し、集計したものが表4である。

表 4 接続表現の出現回数

接続表現	出現回数	接続表現	出現回数	接続表現	出現回数
でも	4,053	それに	37	さて	6
で	3,268	よって	36	だが	6
だから	2,845	ってか	31	では	6
いや	948	てか	31	もしくは	6
だって	893	ところが	30	一方	6
それで	862	したら	27	それでは	5
じゃ	636	いえ	26	ただし	5
じゃあ	590	つまり	21	まずは	5
また	477	あるいは	16	じゃが	4
そしたら	389	すると	15	そうして	4
だけど	293	なのに	15	そもそも	4
しかも	209	ほな	15	てゆーか	4
ほんで	206	ほんなら	14	ついで	3
ただ	200	ほんじゃ	12	でない	3
例えば	187	それなのに	11	なお	3
それから	169	それにしても	11	そうなんと	2
そりゃ	159	ところで	11	もっとも	2
それでも	125	かつ	10	けれども	1
そこで	112	しかし	10	したがって	1
そして	78	そういや	9	それでいて	1
そうすると	76	なら	9	それどころか	1
実は	76	んじゃ	9	なぜなら	1
ないし	53	本当は	9	なので	1
けど	52	こうして	8	ましてや	1
が	50	それなら	8	または	1
そうしたら	47	ちなみに	8	ゆえに	1
それと	41	ですから	8	即	1
それとも	40	追って	8	反面	1

表4の通り、全会話中、接続詞として抽出されたのは、84種類、延べ17,674例であった。原田（2005）においては、小論文のような書き言葉では、「しかし」「また」「そして」「すなわち」「例えば」「つまり」「だが」「あるいは」「したがって」「ところが」等の接続表現が見られ、中でも「しかし」「また」「そして」は突出して出現回数が多いことが明らかになった。それに対して、表4の通り、話し言葉で突出して出現回数の多いものは、「でも」4,053回、「で」3,268回、「だから」2,945回であった。原田（2005）では、小論文のような論理的な文脈展開の書き言葉を対象としたのに対して、本研究の対象は特段に論理構成の求められるようなものではない話し言葉を対象にしたにも関わらず、逆接の意味を表す「しかし」の話し言葉である「でも」が最も多いのは興味深い。同義の「だけど」や「けど」や「が」等を加えると、逆接の意の接続表現の出現回数はさらに増える。しかし、意味用法別に分類すると、逆接よりもむしろ多く見られるのは、順接の確定条件を表す接続表現であることが書き言葉とは異なる。順接の確定条件を表す接続表現に見られたものは「で」「だから」「それで」「ほんで」「そこで」「ですから」等が挙げられる。また、「だって」のように説明・注釈の用法が多く見られることも特徴的であった。

次に、使用されていた接続表現の年代別の属性差に関して、表5に示した。表5における年代別の一人当たりの各接続表現の使用回数は、年代ごとに集計した接続表現それぞれの出現回数を各年代別の延べ参加人数で除して算出しており、1会話における1人の接続表現使用回数を示したものである。

表 5 発話者の年代別一人当たりの各接続詞使用回数 (表 5-1)

	でも	で	だから	いや	だって	それで	じゃ	じゃあ	また	そしたら	だけど	しかも	ほんで	ただ	例えば
10代後半	25.33	10.89	12.11	2.00	5.00	2.78	3.67	3.89	2.33	2.00	0.33	5.78	0.56	0.33	0.22
20代前半	16.40	9.02	7.78	4.29	4.37	1.98	2.90	2.95	1.71	1.11	0.63	1.16	0.40	0.48	0.38
20代後半	22.30	19.72	13.94	4.04	4.39	2.09	2.39	2.98	2.06	1.30	1.43	0.93	1.54	0.83	1.19
30代前半	22.86	16.57	10.38	2.43	3.05	1.71	2.48	2.14	2.00	0.81	0.62	0.52	1.10	1.24	0.95
30代後半	16.40	24.20	13.00	0.80	1.60	3.50	3.00	2.70	1.50	2.40	1.50	0.30	1.70	0.50	0.80
40代前半	8.00	3.60	8.20	0.60	0.80	3.40	1.20	3.20	0.40	0.20	1.00	0.20		0.60	
40代後半	15.33	17.44	14.22	3.50	2.61	8.61	1.28	1.17	2.33	1.94	1.39	0.44	0.67	0.56	1.00
50代前半	6.78	8.78	8.78	2.67	2.00	2.89	1.89	4.22	1.78	0.44	1.78	0.11	1.33	1.22	1.11
50代	9.00	8.00	13.00		2.00	4.00		1.00	2.00		3.00				2.00
50代後半	18.57	15.57	14.71	3.29	4.86	4.14	3.00	0.57	4.43	2.29	0.86	0.29	0.57	0.71	0.43
60代前半	15.94	15.88	20.81	11.06	3.06	6.63	4.75	2.19	2.19	3.06	1.94		0.88	1.19	0.94
60代後半	12.33	10.83	18.50	6.42	7.67	12.42	4.83	1.42	3.00	5.33	3.42	0.58	0.17	2.42	1.75
70代後半	3.00	4.00	14.00	3.00	2.00		5.00	1.00	10.00		7.00	1.00	6.00		
90代前半	5.50	7.75	21.00	3.25	2.00	10.50	0.75	0.75	1.50	5.25	2.75		0.75	3.50	
(平均)	17.62	14.21	12.37	4.12	3.88	3.75	2.77	2.57	2.07	1.69	1.27	0.91	0.90	0.87	0.81

(表 5-2)

	それから	そりゃ	それでも	そこで	そして	そうすると	実は	ないし	けど	が	そうしたら	それと	それとも	それに	よって
10代後半	0.11	0.22	0.33	0.22	0.11	0.33	0.44	1.00	0.22	0.44			0.33		0.33
20代前半	0.21	0.46	0.32	0.25	0.16	0.16	0.27	0.25	0.33	0.14	0.05	0.03	0.11	0.16	0.19
20代後半	0.26	1.17	0.61	0.48	0.31	0.22	0.37	0.28	0.13	0.09	0.30	0.04	0.20	0.17	0.06
30代前半	0.33	0.62	0.38	0.48	0.14	0.48	0.62	0.29	0.05	0.29	0.05	0.05	0.29		
30代後半	0.60	0.20	0.60	1.10		0.10	0.20	0.10	0.10	0.10	0.70		0.30	0.10	0.20
40代前半		0.60	1.40			0.20			0.20	0.40	0.20			0.40	
40代後半	1.61	0.50	0.78	0.89	0.06	0.50	0.44	0.22	0.39	0.17	0.17	0.11	0.17	0.17	0.11
50代前半	0.78	0.22	0.22	0.22	0.44	0.78	0.33		0.11	0.33		0.11	0.11	0.11	0.11
50代	1.00											1.00			
50代後半	0.57		0.43	0.43	0.14	0.14			0.29	0.14		0.86	0.14	0.29	
60代前半	1.56	1.06	0.69	0.19	1.06	0.63	0.06		0.38	0.44	0.25	0.81		0.19	0.31
60代後半	3.58	0.50	0.75	1.50	0.92	0.75	0.58	0.08	0.08	0.50	1.00	0.83	0.33	0.17	0.33
70代後半		11.00	1.00	3.00						1.00					
90代前半	4.75	0.50	2.00	0.50	3.25	0.75	0.25	0.25	0.50	0.50		0.75	0.50	1.00	1.00
(平均)	0.73	0.69	0.54	0.49	0.34	0.33	0.33	0.23	0.23	0.22	0.20	0.18	0.17	0.16	0.16

(表 5-3)

	っつか	てか	ところが	したら	いえ	つまり	あるいは	すると	なのに	ほな	ほんなら	ほんじゃ	それなのに	それにしても	ところで
10代後半	0.22	0.22	0.11		0.22			0.11	0.11				0.11	0.22	
20代前半	0.10	0.27	0.03	0.19	0.05	0.05		0.03	0.13	0.06			0.06	0.03	0.10
20代後半	0.13	0.04	0.06	0.15	0.04	0.07	0.06	0.02	0.07	0.06	0.07	0.13	0.04	0.02	0.02
30代前半	0.24	0.05			0.05			0.05	0.10	0.05	0.38			0.05	0.10
30代後半	0.30	0.10			0.20					0.10			0.10	0.10	
40代前半															
40代後半	0.17	0.33	0.33		0.22		0.33	0.06		0.28	0.06			0.06	0.06
50代前半			0.22					0.11		0.11	0.11	0.11		0.22	0.11
50代															
50代後半			0.14	0.14	0.71			0.14							
60代前半	0.31		0.38	0.06	0.38	0.69	0.06	0.06					0.31	0.06	
60代後半		0.08	0.58	0.33	0.08	0.25	0.50	0.42						0.08	
70代後半								1.00							
90代前半		0.25	0.50	0.25											
(平均)	0.13	0.13	0.13	0.12	0.11	0.09	0.07	0.07	0.07	0.07	0.06	0.05	0.05	0.05	0.05

(表 5-4)

	かつ	しかし	そりや	なら	んじゃ	本当は	こうして	それなら	ちなみに	ですから	追って	さて	だが	では	もしくは
10代後半						0.22		0.11							
20代前半	0.03	0.02	0.10	0.06	0.03	0.05		0.05	0.08			0.05			0.02
20代後半	0.06			0.02	0.07	0.04	0.04	0.02	0.06		0.11	0.04		0.04	0.04
30代前半	0.14							0.05			0.05				0.10
30代後半						0.10		0.10							
40代前半					0.20										
40代後半		0.06	0.17	0.06	0.06		0.06		0.06						
50代前半				0.11			0.44				0.11				
50代															
50代後半						0.14				0.14		0.14		0.14	
60代前半	0.06	0.44								0.19			0.25	0.13	0.06
60代後半	0.08	0.08		0.08	0.08			0.08		0.25			0.17	0.08	
70代後半															
90代前半				0.25			0.25								
(平均)	0.04	0.04	0.04	0.04	0.04	0.04	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03

(表 5-5)

	一方	それでは	ただし	まずは	じゃが	そうして	そもそも	てゆーか	ついで	でない	なお	そうなら	もっとも	けれども	したがって
10代後半						0.11		0.22							
20代前半	0.05	0.02	0.02	0.02	0.05			0.03	0.03	0.02	0.02	0.02			
20代後半		0.04		0.04		0.02	0.04								
30代前半	0.05					0.05			0.05	0.05					
30代後半														0.10	
40代前半							0.40								
40代後半	0.06	0.11													
50代前半					0.11	0.11									
50代										1.00					
50代後半										0.14					0.14
60代前半			0.06	0.13								0.06			
60代後半	0.08		0.25										0.08		
70代後半													1.00		
90代前半										0.25					
(平均)	0.03	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.00

(表 5-6)

	それでいて	それどころか	なぜなら	なので	ましてや	または	ゆえに	即	反面
10代後半									
20代前半						0.02			0.02
20代後半			0.02				0.02	0.02	
30代前半				0.05					
30代後半					0.10				
40代前半									
40代後半	0.06								
50代前半		0.11							
50代									
50代後半									
60代前半									
60代後半									
70代後半									
90代前半									
(平均)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

表5から、全ての年代で使用されている接続表現は、「でも」「で」「だから」「だって」「じゃあ」「また」「だけど」の7種であり、また、表5-1から表5-2に掲載した接続表現は、概ね広い年代で使用されていることが分かる。一方で、表5-3の「ってか」「てか」については、10代後半から30代後半及び40代後半及び60代前半での使用が多く、「なのに」は10代後半から30代前半で使用が見られ、他の年代では使用が見られない。「ほな」は20代前半から30代後半及び40代後半から50代前半で使用が見られ、他の年代では使用されていない。「ほんなら」は20代後半、30代前半、40代後半、50代前半で使用が見られ、他の年代では使用されていない。「ほんじゃ」は20代前半、20代後半で使用が見られ、他の年代では使用がない。「ところで」は20代前半から30代前半及び40代後半から50代前半で使用が見られ、他の年代では使用が見られない。「ちなみに」は20代前半、20代後半で使用が見られ、他の年代では見られない。「ですから」は40代後半及び50代後半から60代後半で使用が見られ、他の年代での使用がない。「だが」は60代前半から60代後半で使用が見られ、他の年代では使用されていない。「では」は20代後半及び50代後半から60代後半で使用が見られ、他の年代では使用がない。「それでは」は20代前半、20代後半、40代後半で使用が見られ、他の年代では見られない。「ただし」は20代前半、60代前半、60代後半で使用が見られ、他の年代では使用が見られない。「そもそも」は20代後半、40代前半で使用が見られ、他の年代での使用がない。「てゆーか」は10代後半と20代前半で使用が見られ、他の年代では使用がない。「もともと」は60代後半、70代後半で使用が見られ、他の年代では使用されていない。

上述のことについて、比較的低い年代と高い年代に分け、視覚化したバブルチャートが図6と図7である。このように、ある特定の接続表現については、使用される発話者の年代が分かれる傾向が高いことが判明した。

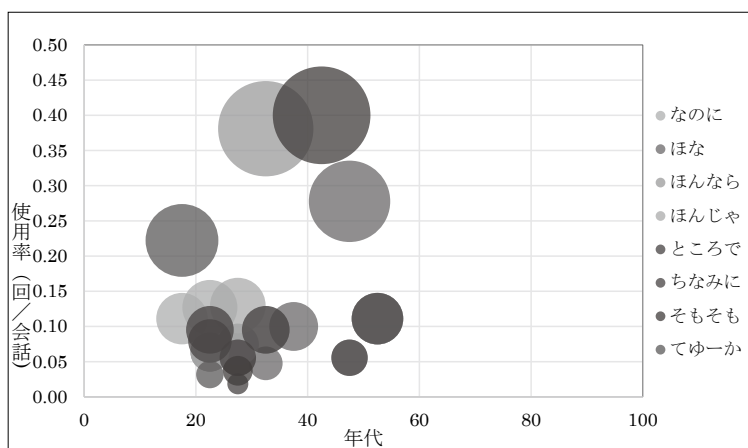


図6 比較的低い年齢で使用された接続表現の年代別の使用率

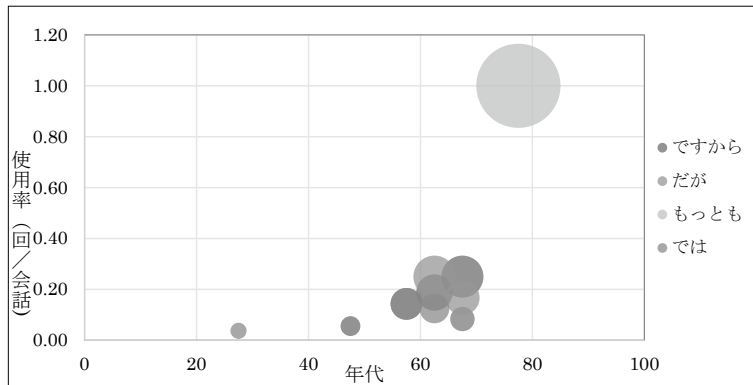


図 7 比較的高い年齢で使用された接続表現の年代別の使用率

3.3.2 目視による接続表現の質的分析

接続表現は、そもそも、その用法から、発話頭に現れるのはもっともなことだが、前項の頻度集計結果において、発話頭出現回数の多い接続表現の上位 5 つである「でも」「で」「だから」「いや」「だって」の前後約 500 文字数を 100 例ずつ無作為に抽出し、目視による分析を行った。その結果、その接続表現が有する主たる意味が無意味化したり、前の発話からの話題の転換や補正を行っていたりするなどして接続表現が使用されており、その接続表現を取り去っても、発話・談話全体の意味や内容が変わらず、いわばフィルター化しているのではないかと思われるものが確認された。

最も出現回数の多かった「でも」については、例文 (1) のように逆接・確定条件として使用されているもの以外に、例文 (2) や例文 (3) のように、逆接の意味は失われて、前の発話から話題の転換や補正を試みようとしている例が約 1 割見られた。

(以下の例文中の英数字は発話者番号を表す。)

例文 (1)

M015: なんで? テレビなんか見ちゃいかんて。勉強しなさいって。

F004: うちね、なんかね。すごくちっちゃいころのね、しつけは厳しかった。うん。でも そのあとは、すごい、なーなーだね。おもう。

M034: すごくちっちゃいころの、しつけが厳しかった?

例文 (2)

F004: 違う違う違う。切ると同時に焼けるんだって。あ、やばい。お父さん? 何でお父さんまだこんな時間に起きてんの? やだー。邪魔されないように、しないと。でも さ、こんなさ、あのとき思ったんだけど、思ったんですけどねえ、ちょっとちょっと、思ったんですけどねえ、声が似てるじゃん。特に F018 ちゃんと F004 なんて。だ

からさあ、文字化するときすごい大変だと思う。

F018：これだれが文字化すんの？

F004：業者。

例文 (3)

F005：ロシアもヨーロッパなんです。で、昨日もらったのは、も、ね、直したのをもらったんです。

F067：直してあるのを。

F004：ふーん、そっか、そっか。うーん。そう、でも何かわりと最近のテレビ、あ、最近のビデオはどっちもオッケーだとかっていうのもあるらしいんだけどー。うん。でも、確か、さん、違ったかなー、うーん、うん、そう、どっちもオッケーなテレビ、ビデオがあるとかいう話を聞いたんだけど。うん。

「で」については、例文 (4) のように順接・確定条件として用いられているもの以外に、例文 (5) のように無意味化し、その談話の文脈に必要がないと思われるものが約 1 割見られた。

例文 (4)

F001：なんかさ、失敗するとすぐ周りのせいとか人のせいにしてさあ。

M033：いいじゃん。

F001：で、そういうふうには逃げ道を用意している人ってたいがい失敗するよね。

M033：いや、僕は大丈夫だよ。適当にやるから。

例文 (5)

M033：どこだよ。

F001：こないだねえ、なんか先生が話してたんだけどー、なんかね、チョウの研究をしている人がいるんだって。チョウチョのね。で、なんかチョウの。

M033：チョウチョウって、町の町長とかそういうの？

同じく、順接・確定条件の「だから」については、発話頭出現回数が多かったものの、フィルター化したような例は見られなかった。しかし、書き言葉においては、前に述べたことが後に述べたことの原因になる際に用いられることが多いのに対して、例文 (6) のように「繰り返し言っているように」という意味で用いられているものが約 3 割も確認できた。

例文 (6)

F001: だから、お母さんの時代でさえ 27 なんだからー。あたしの時代だったら、30 はあたりまえ。

F079: まさか、そりゃあ、F001 そんな遅くだめ。

F001: 遅、だめってそんな。だめって言われても困るけど。

(中略)

F001: だから今は 30 歳超えて子どもを産むのは。

F079: よかないよ、あんな遅遅で。

F001: だから、30 過ぎて子ども産むのは今、普通だつてば。だつて T ちゃんだつてこないだ、結婚したの 32 でしょう？

「いや」については、形態素解析で「感動詞」や「接続詞」として抽出されるものが見受けられたが、そのうち「接続詞」として抽出されたものを 100 例見ると、実際は例文 (7) や例文 (8) のように、話し手が相手の発話を打ち消して話を切り出すものが約 8 割見られ、例文 (9) のように、本来「うん」と肯定して答えても良さそうなところを話し手の何らかの心的態度の表出で「いや」を使用している例や、例文 (10) のように、話し手が自らの発話を補正しながら使用しているか、もしくは無意味化したようなフィラーとして使用していると思われるものが約 2 割見られ、「いや」が「接続詞」として抽出されたのがもはや不可解である。

例文 (7)

M034: 弟さんもさ、F 大なんだって？

F014: いや、違いますよ、うちは。

例文 (8)

F005: いーえ、全然。私はどっちかというと待っててー、えー、何か、何かちょっと果物食べたくないーっていうタイプで。

F034: そう？ 率先して切るタイプだと思っていたのに。

F005: いや、うちではそんなことないですよ。

F113: あ、そう。

例文 (9)

F019: 独学。1 か月であきらめた。

F004: そうなんだ。え、なんか目的あったの？

F019: いや、院の入試を考えてたときに、なんかヨーロッパ言語を 1 つ知らない。韓国

語では入らない、入れない、受験できないって聞いていたから。

F004：ふーん、そうか。

例文 (10)

F005：でも、私、おつまみに、食べるのが、ごま豆腐食べるのが好きなんですよー。

F034：おつまみとか言ってるし。

F113：まあ、まあ、あの、大体おいしいものっていうの、あの、お酒に合うんですよ。そういうのね。

F005：大好きなんですよ。いや、ごま豆腐もやっぱにおいがするから私すごい好きなんですよ。

F034：はー、香りの強いものが。

F005：うーん。結構好きで。

「だって」は、例文 (11) のように、注釈・説明を表す接続表現がほとんどだったが、例文 (12) のように直接的に注釈・説明を表している部分に「だって」が使用されているのではなく、その発話部分においては「だって」は不必要で、無意味化したフィラーのごとく、使用されているような例が約 1 割程度確認できた。

例文 (11)

F037：(略) ドコモって高くなーい？

F067：高い。うん。

F037：私ねー。

F067：Jフォンかどっかにするのー？

F037：ううん、au。うん。

F067：うん。留学生の人に聞いたら au が安いって言った。

F037：だって半額でしょう、学生だから。(以下、略)

例文 (12)

F021：何か、中退しちゃおうっていう気はないみたいだけど。何とかなるさっていう。遅いんだけどね。そういう考えだと私もいつまでも待っててくれると思ってると思うよってことを言いたかったわけ、いいかげんにしないと、ほんと見捨てるからねって。

F155：うん。そっかー。ちょっと、ちょっと自立してないねーって感じよね。

F021：あー、ちょっと。だからさ、ほんとうらやましい、ちゃんと働いててさー、お金も取れててさー、ちゃんとしてる彼氏で、それでけんかするなら、すごいけんかしたのよ、私は。とっってもけんかしたい、そんなこと意見くいちがっても。でも、何に

も持ってないゼロの人と、まともにやり合うこと自体があほらしくて、知らんって。もう。

F155：そう。何言ったって、だって絶対、おかしい。

F021：だって、開き直ってるの。何もないことは知ってるけど、自分自身で。だってしょうがないじゃんって。それはないじゃないでしょうって言って。すごい言い訳がましいのね。何か、すごい考えるんだって、これじゃまずい、これじゃまずいって。

F155：何でよーって感じなの。

3.3.3 接続表現に関する考察

頻度集計による分析の結果、発話頭に現れる接続表現について、以下のようにまとめられる。

- i) 「でも」「で」「だから」「だって」「じゃあ」「また」「だけど」などは、出現回数が多く、また、全ての年代で使用されている。
- ii) 「なのに」「ほな」「ほんなら」「ほんじゃ」「ところで」「ちなみに」「そもそも」「てゆーか」などのように、比較的低い年代で多く使用され、且つ、比較的高い年代では使用されにくい接続表現がある。
- iii) 「ですから」「だが」「もっとも」「では」などについては、比較的高い年代で多く使用され、且つ、比較的低い年代では使用回数が非常に少ない。

このように、ある特定の接続表現については、使用される発話者の年代が分かれる傾向が高いことが判明した。

目視による分析で得られた考察結果については、以下のようにまとめる。

- i) 「でも」「で」「だって」のように、各々、逆接・確定、注釈・説明、順接・確定などの意味・用法が失われ、無意味化し、フィルター化された接続表現がある。
- ii) 「だから」のように、話し言葉の中で多用されていても、フィルター化しにくい接続表現もある。
- iii) 形態素解析の結果、「接続詞」ではないと思われるにもかかわらず、「いや」のように「接続詞」として抽出されてしまうものがある。

3.4 副詞の分析

3.4.1 テキストマイニングによる副詞の分析

会話の中で、発話頭や発話中や発話末に見られる談話標識のうち、フィルターや接続表現に関して分析を行ってきたが、発話頭に現れる副詞が思いのほか多いことが確認され、それらにも何らかの特徴が見られると思われたため、本項では、副詞についても接続表現と同様に、頻度集計や目視による分析を試みた。

全会話中で、形態素解析ツール MeCab により、副詞として抽出されたものは、823種類、延べ42,221例だった。出現回数の多い順に、「もう」「そう」「ちょっと」「こう」「まあ」

「どう」「やっぱり」「全然」「結構」「ほんとに」「たぶん」「まだ」「あんまり」「よく」「別に」「なんで／何で」「どー」「ちゃんと」「やっぱ」「いろいろ」「そんなに」「いっぱい」「一応」「すぐ」「ずっと」「もっと」「いつも」「きっと」等が見られた。なお、接続表現は発話頭に現れるものが、そもそも非常に多かったため、発話頭出現率（＝発話頭出現回数／出現回数）を算出したのは、出現回数上位5つに限定したが、副詞については、副詞として抽出されたもの全てに関して、発話頭出現率を算出し、その結果、発話頭出現率が10%以上のもの28種類、延べ3,396例を表6に示した。

表6 副詞の出現回数と発話頭出現率

	副詞出現回数合計	発話頭出現回数合計	発話頭出現率
やっぱ	317	284	89.6
なんで／何で	586	317	54.1
まあ	1102	396	35.9
かなり	127	44	34.6
まず	104	36	34.6
たぶん	630	171	27.1
ほんとに	649	172	26.5
もう	3383	879	26.0
本当に	163	42	25.8
別に	411	98	23.8
よく	492	117	23.8
ずーっと	111	26	23.4
結構	803	184	22.9
ちょうど	128	28	21.9
もっと	231	47	20.3
もちろん	123	24	19.5
とにかく	138	26	18.8
いつも	210	39	18.6
ほとんど	117	21	17.9
やっぱり	1078	193	17.9
そんなに	293	51	17.4
きっと	188	32	17.0
なんて	161	27	16.8
初めて	121	19	15.7
ちゃんと	373	54	14.5
何とか	114	16	14.0
いっぱい	242	28	11.6
ずっと	232	25	10.8

表6から、発話頭出現率が高かったものは、高い順に「やっぱ」「なんで／何で」「まあ」「かなり」「まず」と続いている。

次に、使用されていた副詞のうち、出現回数が100以上のもの51種類についての発話者の年代別の使用状況を表7に示した。表7において、発話者の年代別の一人当たりの各副詞の使用回数は、発話者の年代ごとに集計した副詞それぞれの出現回数を各年代別の延べ参加人数で除して算出しており、1会話における1人の副詞使用回数を

示したものである。なお、副詞は接続表現と比べ、種類も出現回数も極めて多いため、表7では、それぞれの副詞の使用回数が平均使用回数150%以上のものを特に使用の多いものとし網掛けで表示し、視覚化した。また、40代後半と50代前半を境に比較的年代が低いグループと比較的年代の高いグループに分類し、それぞれの副詞の年代別の属性差を分析した。

比較的年代が低いグループで多用されているのは、「そう」「こう」「まあ」「結構」「たぶん」「別に」「どー」「やっぱ」「きっと」などである。一方、比較的年代が高いグループで多用されているのは、「もう」「ちょっと」「やっぱり」「まだ」「あんまり」「よく」「ちゃんと」「いろいろ」「ずっと」「もっと」「ちょうど」「とにかく」「何とか」「また」「ああ」「どうして」などであった。

表7 発話者の年代別一人当たりの各副詞使用回数(表7-1)

	そう	もう	ちょっと	こう	まあ	どう	やっぱ	全然	結構	ほんとに	たぶん	まだ	あんまり	よく	別に
10代後半	83.22	18.11	10.78	5.33	3.22	7.89	1.89	5.22	8.33	5.33	6.22	3.67	1.78	2.33	3.22
20代前半	34.44	14.52	6.81	6.05	4.35	6.78	2.87	4.54	2.97	3.48	3.35	2.57	1.71	2.05	2.48
20代後半	52.22	17.28	10.72	13.74	9.09	6.91	7.61	5.63	6.17	3.26	5.00	3.59	3.44	3.00	3.61
30代前半	57.05	15.76	8.90	6.57	5.90	5.48	6.48	4.00	4.76	4.14	4.10	2.95	2.76	3.24	2.10
30代後半	32.40	11.10	8.90	4.80	3.70	6.60	3.10	4.20	4.70	1.30	1.70	3.60	2.30	1.90	1.50
40代前半	22.80	6.40	3.60	3.20	3.40	3.40	3.80	2.00	3.80	3.60	0.80	2.20	0.80	2.00	1.00
40代後半	37.06	21.28	10.22	9.44	9.33	3.56	7.11	3.00	3.61	3.72	1.44	3.28	3.17	2.94	2.17
50代前半	18.00	10.78	9.44	6.22	2.56	2.44	4.44	2.89	1.00	0.44	0.78	2.67	1.44	1.11	0.67
50代	18.00	11.00	16.00	3.00	1.00	3.00	1.00	4.00		6.00	1.00	2.00		2.00	
50代後半	38.71	19.29	8.71	6.86	3.71	3.57	5.00	2.43	5.43	4.00	1.43	2.14	2.71	1.57	0.29
60代前半	56.44	30.13	7.50	4.63	7.44	3.63	10.06	3.50	2.69	4.00	1.94	4.81	2.25	2.63	0.81
60代後半	51.83	25.83	10.58	5.75	5.42	5.92	7.08	3.83	2.17	2.75	1.75	3.00	3.83	2.50	0.83
70代後半	38.00	9.00	3.00	8.00	4.00	8.00	1.00	3.00	1.00			2.00	13.00	9.00	1.00
90代前半	19.00	20.00	8.50	8.25	4.50	2.75	3.00	1.75	0.50	2.00		2.00	3.00	2.75	
平均使用率①	44.00	17.30	8.80	7.90	6.00	5.70	5.40	4.20	4.10	3.30	3.20	3.10	2.50	2.50	2.20
①の150%	66.00	25.90	13.20	11.80	9.00	8.50	8.10	6.30	6.10	4.90	4.80	4.60	3.70	3.70	3.30

(表7-2)

	なんで	どー	ちゃんと	やっぱ	いろいろ	さ	そんなに	いっぱい	一応	すぐ	ずっと	もっと	いつも	何で	きっと
10代後半	2.56	3.11	2.67	2.89	0.44	4.56	1.00	1.33	0.78	1.89	0.78	1.44	2.11	1.44	0.33
20代前半	4.02	2.56	1.46	2.27	1.14	1.79	1.08	1.41	1.11	0.86	1.25	0.89	0.79	1.29	0.70
20代後半	2.43	3.61	2.28	1.78	1.85	2.67	2.13	1.11	1.83	1.15	1.19	1.39	0.78	1.11	1.39
30代前半	1.90	2.76	2.67	1.90	1.67	1.71	2.24	1.76	1.71	1.24	1.19	1.57	1.33	0.24	1.33
30代後半	0.40	1.20	1.50	0.40	1.00	0.50	1.30	2.00	1.60	1.10	0.90	1.10	0.90	0.60	0.90
40代前半	0.60	0.20	0.60	0.20	1.20	0.20	2.20	0.40	0.60	0.80	0.60	0.40	0.80		1.40
40代後半	0.50	0.83	1.00	1.56	2.00	0.28	1.72	1.44	0.94	1.39	1.78	1.33	0.67	1.78	1.11
50代前半	0.22	0.89	1.89	0.56	1.78		0.22	0.56	1.56	1.56	0.67	1.00	0.56	0.22	
50代			1.00		1.00		1.00				1.00	1.00	1.00		
50代後半	1.14	1.43	1.43	1.29	1.43	0.43	1.86	3.57	0.43	1.71	0.57	1.00	1.14	0.71	0.86
60代前半	0.94		3.81	1.50	3.31	0.38	1.25	0.81	1.38	1.69	2.00	1.81	1.88	0.31	0.50
60代後半	0.92	0.17	2.42	0.25	1.58	0.58	0.92	1.42	1.00	2.75	1.25	0.83	1.42	0.67	0.83
70代後半	9.00		3.00		8.00	1.00	8.00	1.00		5.00	1.00				1.00
90代前半		0.75	2.75		1.50	0.25	1.25	1.00		1.75	1.00	0.25	4.50	0.25	0.25
平均使用率①	2.20	2.10	2.00	1.60	1.60	1.50	1.50	1.30	1.30	1.20	1.20	1.10	1.00	0.90	0.90
①の150%	3.30	3.10	3.00	2.40	2.40	2.20	2.20	1.90	1.90	1.80	1.80	1.60	1.50	1.30	1.30

(表 7-3)

	本当に	なんて	ちょうど	もちろん	とにかく	なかなか	ほとんど	かなり	何とか	とりあえず	初めて	また	ずーっと	まず	ああ
10代後半	3.78	0.44	0.78	0.11	0.11	1.00	0.22	1.67	0.11	0.44	0.44	0.56	0.22	0.44	0.33
20代前半	0.94	0.52	0.48	0.35	0.37	0.49	0.54	0.56	0.46	0.76	0.33	0.54	0.51	0.59	0.29
20代後半	0.70	0.98	0.93	0.91	0.72	0.81	0.56	0.93	0.48	0.74	1.00	0.56	0.67	0.48	0.48
30代前半	0.52	2.67	0.43	0.71	0.86	1.10	0.62	0.24	0.38	0.67	0.62	0.57	0.90	0.71	0.57
30代後半	0.30	0.60	0.50	0.30	0.70	0.40	0.70	0.10	0.30	0.50	0.90	0.30		0.50	0.70
40代前半	0.40	0.40	0.20	0.40	0.60			0.20	0.80	0.20	0.60		0.40	0.40	0.20
40代後半	0.22	1.11	0.89	0.61	0.56	0.72	1.00	0.67	0.28	0.67	0.39	0.50	0.50	0.61	0.50
50代前半	0.11	0.44	0.11	0.22	1.89	0.44	0.44		0.67	0.33	0.56	0.67	0.22	0.78	0.22
50代			1.00	2.00				2.00	1.00		3.00		1.00		3.00
50代後半	2.71	0.29	1.43	0.57	0.29	0.57	0.71	0.29	0.57	0.14	0.14	0.43	0.86	0.14	1.14
60代前半	0.56	0.63	0.63	2.25	2.19	0.94	1.31	0.81	0.94	0.63	0.69	0.94	0.19	0.75	0.75
60代後半	1.33	0.17	1.17	1.08	0.17	0.75	1.33	0.58	2.67	0.08	0.58	1.00	1.08	1.00	0.92
70代後半		2.00						1.00				1.00			9.00
90代前半	0.50		2.00		0.25	0.25	0.75		2.25		0.25	1.25	2.25		0.25
平均使用率①	0.80	0.80	0.70	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.50	0.50	0.50	0.50
①の150%	1.20	1.20	1.00	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.70	0.70	0.70	0.70

(表 7-4)

	どうして	少し	いやいや	実際	そのまま	だんだん
10代後半	0.11	0.11	0.11	1.33	0.56	0.33
20代前半	0.17	0.22	0.37	0.35	0.35	0.16
20代後半	0.57	0.52	0.33	0.48	0.41	0.24
30代前半	0.29	0.14	0.62	0.90	0.52	0.43
30代後半	0.30	0.30		0.20	0.40	1.00
40代前半	0.20	0.80	0.20	0.40	0.20	0.40
40代後半	0.06	0.39	1.00	0.22	0.28	0.56
50代前半	0.44	0.33	0.11		0.67	0.44
50代	1.00	3.00		2.00		
50代後半	1.00		0.86		0.57	1.29
60代前半	1.13	1.25	0.38	0.63	0.38	1.56
60代後半	1.75	0.83	1.17	0.17	0.17	0.33
70代後半	1.00	1.00				
90代前半	0.50	1.50			3.00	0.25
平均使用率①	0.40	0.40	0.40	0.40	0.40	0.40
①の150%	0.60	0.60	0.60	0.60	0.60	0.60

3.4.2 目視による副詞の質的分析

副詞は、文の中で主に用言に係るのだが、話し言葉においては、ある一定の副詞が発話頭に多く現れて用いられることも多い。ここでは、全体の出現回数が多く、且つ、発話頭出現回数の多い「もう」「まあ」⁴「やっぱ」「やっぱり」「なんで／何で」「結構」を取り挙げ、それらの前後約 300 文字数を 100 例ずつ無作為に抽出し、目視による分析を行った。

その結果、「もう」は、例文 (13) のように、時間や程度がある基準を超えたり、ある基準に近づいたりする意味を表す副詞や、既にある基準や状態に達しているのに更に加える意味を表す副詞等が約 6 割見られたのに対して、例文 (14) や例文 (15) のように、上述の意味はなく、フィルターと言っても良いと思われるものが約 4 割見られた。(以下の例文中の英数字は発話者番号を表す。)

例文 (13)

F137: うん、そう。地元の友達はいいいね、ほんと。私もね、今度友達が就職決定して、新宿かどっかで飲もっかって。い、うん、友達と。

F114: え、タメの友達?

F137: うん、そう。

F114: ふーん。もう決まったんだ。

F137: うん。短大行って。で、たぶんお父さんの会社に就職するとか言って。(以下、略)

例文 (14)

F086: そうなん。すごいなー、もうあれ。ほんっとにやりたくない。助手も助手だしさ。

F046: 助手、本当嫌だよー、あの人ー。

F086: Tさんでしょ。

F046: Tさんだよ。

F086: もう、なんなーん、あの人。キモいんさ。とか言ってこれ入っててさ、まあ、別にいい。全然何も知らない。あの人別に先生じゃないよね。

F046: うん。大丈夫や。

例文 (15)

F073: えっ? おいしいの?

F092: おいしくないけど、体脂肪が燃焼されるからって言って。

F073: そーら飲まなあかん。

F092: 飲まないかんね。

F073: もうあれやもん。なんか、なんかさ。大丈夫?

F092: 大丈夫だよ。えっ?

「まあ」については、例えば、驚いたり感心したりする時に発せられる用例ではなく、例文 (16) のように相づちや、例文 (17) のように、その場の発話者「M034」と「M015」に配慮し、その場を収めるはたらきや、例文 (18) のように、間つなぎのはたらきなどをし、フィラー化しているものが約9割を占めていた。

例文 (16)

M006: 卒論とかは?

F122: 卒論はね、うちの大学は、なかったんですよー。

M006: ああ、そうなんですか。

F122: まあゼミ論文だけで。だからそんな大したことはなかったですね。

M006: ふーん。うん。

例文 (17)

M015: そんなことないでしょ。魅力はあるんだから。

M034: そんなことはない。

M015: そんなことはないでしょ。考えすぎでしょ、それは。

F004: まあな。まあな。

M034: キャラが変わった。

例文 (18)

F005: うん。メール、今回は、ほとんど出さなかった。なんか、すごい忙しくて。もう、いっぱいいっぱい暮らしてた。

F004: あ、ほんとに一。そうか、そうか。まあ、私もいっぱいいっぱいだったんだけど、だけど、うん、まあ、たまにって、ほんとに、えー、4か月や6か月にいっぺんぐらいなんだけども。

「やっぱ」「やっぱり」は、発話者が事前に予想した通りだったと確認する意味や、話し手自身の判断の正しさを聞き手に伝えるなど、話し手の評価の客観性の高さを表したいという意図で使用されているものがほとんどであったが、例文 (19) や例文 (20) のように、その評価の意味も失われ、フィルター化しているのではないかと思われるものが、少数見られた。

例文 (19)

F093: うん。絵手紙は別に。

F097: 絵手紙はね。

F093: 飾っとかんでも。

F097: うん、そうそう。出せばいいの。あげればいいの。お便りを。やっぱりね、あの、顔彩って言って、あの、絵の具もそういうのあるの。

F093: 顔彩?

例文 (20)

F034: え、何、何が得意料理ですか。

F113: 何かって、いえ、精進料理みたいなもんばかり食べてますよ。もう。うん。

F034: え、和食がお好き。

F113: 和食ですね、ほとんど。うん。あんまり一、あの。でも、あの、鶏は好きなんですけどもううちではその、から揚げとかそういうものは食べなくて、ここのおいしいなと思ってときどき食べたくなる。あと、焼き魚とか。やっぱね、でもね、結構うるさいんですよ。その、おだしとかね。私が。主人なんか何食べてもわかん

ないけど。何か、おだしをとって、それからだからね、結構時間かかります。

F034：そうですね。

「なんで／何で」については、原因・理由などを尋ねる意味はもとより、例文(21)のように、聞き手に対する反論の意味を表すものが2割弱、例文(22)のように話し手自身が自問自答する例が少数見られたが、副詞として抽出された「なんで／何で」に関してはフィラー化していると思われるものは見受けられなかった。

例文(21)

M033：ていうか、お前、恥ずかしいよ。しらたきを切らないで、なべに入れてしまう君は恥ずかしい。

F001：恥ずかしくないよ、全然。

M033：恥ずかしい。

F001：恥ずかしくない。

M033：恥ずかしい。

F001：全然恥ずかしくない。

M033：恥ずかしいね。あれだよ、もう、トイレに入ってドアを開けたまま、うんこするよなもんだよ。

F001：なんでだよ。

M033：恥ずかしい。

F001：恥ずかしくないよ。

例文(22)

F046：テスト？あー。

F086：宿題やってないんだ、まだ全然。あーあ、明日は、明日は忙しい。本当に、なんでこんなに忙しいんだ。あたしね、火、水って本当はさ、最後、授業最後までだからさー、もうこれで帰って、寮に帰って、ゆっくりっていう日なのにさ、両方とも予定が入るの何気に痛い。

F046：ねー、痛いよねー。

完全ではないが、それなりに満足がいくという話し手の評価の意味が含まれる「結構」については、例文(23)のように、主に場つなぎや間つなぎとしてだけではなく、例文(24)のように発話権の維持と間つなぎを兼ねた機能でフィラー化しているものなどが1割弱見られた。

例文 (23)

- M017: 会わない。そうそうそう。あ、あの先生、違う？ どうも、ごぶさたしております。
M029: どうもごぶさたしております。なんか相当歩き回りました。いやね、東急インだと
思ったんですけどね。
F098: ええ。
M017: ちょっと、ちょっといすをもらいましょう。
F098: うん。あ、すみません。
M017: 今日、広島からお着きになって。
F098: 結構、あれなんですよ。
M029: おっちょこちよいなんです。
F098: 前、ガーランドも。
M029: ガーランドをガーデンホテルとまちがえて。

例文 (24)

- F004: 何しゃべろう。何しゃべろう。あっ、もうわかってんの。男の子、女の子かって。
F091: わからないの、まだ。
F004: えっ、まだわかんないの？
F091: うん。
F004: えっ？
F091: ちょうど検診に行ったときに向きとかが見えないところにいるみたいでわかんない
みたい。うん。そうなの。
F004: うーん。結構あれだねー。きましたねー。へえー。うん。へーえ。えーっ、でも変
な感じじゃない？ 自分のお腹がふくらんでくるのって。
F091: うーん、怖い。

3.4.3 副詞に関する考察

頻度集計による分析の結果、発話頭に現れる副詞として抽出されたものについては、以下のようにまとめた。

- i) 「そう」「こう」「まあ」「結構」「たぶん」「別に」「どー」「やっぱ」「きっと」などのように、比較的低い年代で多用される副詞がある。
- ii) 「もう」「ちょっと」「やっぱり」「まだ」「あんまり」「よく」「ちゃんと」「いろいろ」「ずっと」「もっと」「ちょうど」「とにかく」「何とか」「また」「ああ」「どうして」などのように、比較的高い年代で多用される副詞がある。

目視による質的分析の結果、フィルター化する割合の傾向から、話し言葉における副詞を主に以下の4つに分類した。

- i) 「まあ」や「もう」など感動詞的なものは、副詞として抽出されたと言えども、

副詞の意味・用法に代わり、間つなぎ、相づち、他の発話者に配慮し、その場を収めるはたらきをするなど、フィラー化しやすい。

- ii) 話し手の客観的な評価を表す「やっぱ」「やっぱり」のような副詞の中に、評価性が失われ、フィラー化しているものが少数見られる。
- iii) 話し手のより主観的な評価としてはたらく「結構」に、発話権の維持、間つなぎ、場つなぎとしてはたらき、フィラー化するものが1割弱見られる。
- iv) 「なんで／何で」のように発話者の疑問、反論、自問自答などの心的態度を強く表すが、フィラー化しにくい副詞も勿論ある。

以上のことから、全ての副詞がフィラー化するわけではないが、「まあ」や「もう」のように限りなくフィラーに近いものがある一方で、客観的側面・主観的側面という評価性に関しては両極にある「やっぱり」と「結構」にフィラー化する傾向が出現しつつあることが明らかになった。

4 おわりに

本稿では、フィラー、接続表現、副詞といった談話標識に着目し、話し言葉の中で起きる現象について、テキストマイニングを用いた統計的手法と目視による分析により、考察を行った。

フィラーはもっぱら聞き手や会話の場そのものを意識して用いられ、接続表現は談話の内容に関わり、副詞は主に用言に係って文中に出現する。しかし、談話内容に関わる意味・用法が弱まり、無意味と化した接続表現が出現し、また、用言から離れ、発話頭出現率の高い副詞の中に、シンタクスレベルの機能が弱まったものが存在することが明らかになった。それらを本稿では、接続表現や副詞のフィラー化として取り挙げた。

接続表現に言及すると、「でも」「で」「だって」のような接続表現について、本来の意味・用法が失われ、フィラー化しているものが一部存在する一方で、「だから」のように、話し言葉の中で多用されていても、フィラー化しにくい接続表現もあることが判明した。

そして、副詞については、「まあ」や「もう」などはフィラー化しやすく、「やっぱり」「結構」などはフィラー化しやすいとまではいかないが、フィラー化するものを見出した。

また、テキストマイニングによる統計的手法では、ある年代で頻繁に使用される傾向のあるフィラーや接続表現や副詞が明らかになった。たしかに、比較的低い年代による発話において、接続表現や副詞の元来の意味・機能性が低い使用傾向があり、使用する表現に偏りが見られるものの、フィラーやフィラー化しやすい接続表現や副詞を使用するのは比較的低い年代に限らず、比較的高い年代でも使用されているものもあった。

現在、他の接続表現や副詞に関しても、フィラー化しやすいもの、しにくいものと、

接続詞や副詞の分類との関連性を明示するために、より詳細な分析を別稿にて行っている。さらに、形態素解析ツールを用いた言語の統計的研究を行うには、対象とする語や表現について、形態素解析ツールの辞書機能への登録等、形態素解析の精度を上げることも避けては通れない、今後の課題である。

注

- 1 『名大会話コーパス』は、科学研究費基盤研究 (B) (2) 「日本語学習辞書編纂に向けた電子化コーパス利用によるコロケーション研究」(平成 13 年度～15 年度 研究代表者 大曾美恵子) の一環として作成された、日本語母語話者同士の雑談を文字化したコーパスである。現在は国立国語研究所に移管され、文字化テキストなどを公開している。
- 2 Rとは、ニュージーランドのオークランド大学統計学科の Ross Ihaka とアメリカのハーバード大学の生物統計学科の Robert Gentleman により開発が始められ、1997 年以降、多くの賛同者によって開発が続けられているオープンソース方式のデータ解析・処理の専用ソフトのことである。詳しくは、金明哲 (2017) 『R によるデータサイエンス』を参照されたい。
- 3 本稿では「会話」を発話のまとまりとして捉え、「発話」という言葉と区別して使用している。
- 4 形態素解析の結果、「まあ」は「フィラー」や「感動詞」として抽出されるものもある。本稿の頻度集計表や目視による分析における「まあ」は、形態素解析で「副詞」として抽出されたものを対象としている。

参考文献

- 金明哲 (2017) 『R によるデータサイエンス』第 2 版, 森北出版
- 定延利之 (2016) 『コミュニケーションへの言語的接近』ひつじ書房
- 泉子・K・メイナード (1997) 『談話分析の可能性』くろしお出版
- 中島悦子 (2011) 『自然談話の文法—疑問表現・応答詞・あいづち・フィラー・無助詞—』おうふう
- 原田朋子 (2002) 「中国語母語学生の作文における文脈展開上の問題点—接続表現の観点から—」『平成 14 年度日本語教育学会第 7 回研究集会予稿集』pp.45-48.
- (2005) 「接続表現から見た文脈展開—日本語母語話者と上級日本語学習者の小論文比較—」『同志社女子大学大学院 文学研究科紀要』第 5 号, pp.103-120.
- (2019) 「日本語母語話者と上級日本語学習者の小論文の比較—テキストマイニング手法と目視による分析を通して—」『同志社大学 日本語・日本文化研究』第 16 号, pp.1-15.
- 山根智恵 (2002) 『日本語の談話におけるフィラー』くろしお出版
- 渡辺実 (1996) 『日本語概説』岩波書店

