

# 日本語における子どものフィラー使用の発達的变化 — 『子ども版日本語日常会話コーパス』に基づく分析 —

田中 真理子 居關 友里子 小磯 花絵

国立国語研究所

{tmariko,iseki,koiso}@ninjal.ac.jp

## 概要

本研究は『子ども版日本語日常会話コーパス』を活用し、1歳から7歳までの子どもを対象に、フィラーの形式、出現位置、発話単位長に着目して子どものフィラー使用の発達的变化を分析した。その結果、フィラー使用は月齢に伴い一様に変化するのではなく、特定の時期を境に変化し、その様相はフィラー形式や出現位置によって異なることが示された。また長い発話では、24~60か月にかけてフィラー出現率が緩やかに上昇した後、60か月以降に急速に低下するという発達の推移が確認された。

## 1 はじめに

会話において頻繁に用いられる「あの」や「えーと」といったフィラーは、命題内容に直接関与しない余剰成分として片付けられるものではなく、発話の構築や相互行為の調整に関わる重要な発話要素として捉えられてきた。このような観点から、フィラーについては言語学をはじめとするさまざまな研究領域において、数多くの研究が蓄積されてきた[1, 2]。一方、子どものフィラー使用を扱った研究としては、複数のフィラー形式を子どもも含め観察した研究[3]、CHILDSや『日本語日常会話コーパス』を用いて子どものフィラー使用を定量的・定性的に分析した研究[4, 5]、また幼児の発話に生じる非流暢性とその変化に着目する中でフィラーを扱った研究[6]などがあるが、その数は限られており、また扱っているデータ規模も必ずしも大きいとは言えない。

そこで本研究では、『子ども版日本語日常会話コーパス』を中心的資料とし、『日本語日常会話コーパス』に含まれる成人データも比較資料として活用することにより、子どもの成長に伴うフィラー使用の変化の実態を定量的に明らかにすることを目的とする。

## 2 方法

### 2.1 データ

分析には、国立国語研究所で現在構築中の『子ども版日本語日常会話コーパス』(CEJC-Child)を用いた。CEJC-Childは、子どもの日常会話を経年的に記録した100時間規模の会話コーパスであり、そのうち50時間の会話がモニター版として2025年3月に公開されている[7]。本研究では、現在整備中のデータを含む100時間分の会話から、調査対象児12名のデータを分析対象とした(付録・表3)。このうち、各年齢区分において異なり話者数が4名以上確保できる年齢帯として、1歳(1;0-1;11)から7歳(7;0-7;11)を対象とした。また比較対象として『日本語日常会話コーパス』(CEJC)[8]も用いた。雑談中心のCEJC-Childに合わせ、CEJCも雑談に限定した上で、20歳以上の成人話者のデータを抽出した(各年齢区分のデータ規模は付録・表4参照)。

### 2.2 対象とするフィラー

本研究におけるフィラーの認定は、CEJCで採用した基準に準拠した[9]。具体的には、CEJCで定義されている以下の基本要素、およびそれに一定の派生要素が後続する形式をフィラーとして扱った。

基本要素：

あ(ー)、い(ー)、う(ー)、え(ー)、お(ー)、ん(ー)、  
と(ー)\*、つと(ー)\*、  
あ(ー)(ん)(ー)の(ー)\*、そ(ー)(ん)(ー)の(ー)\*、  
う(ー)(ん)(ー)つと(ー)\*、ん(ー)つと(ー)\*、  
あ(ー)(ん)(ー)つと(ー)\*、え(ー)つと(ー)\*

派生要素 (0回あるいは1回後続する形式)：

ですね(ー)、っすね(ー)  
ね(ー)、さ(ー) (\*を付した表現に後続する場合のみ)

分析にあたっては、CEJC-Childに付与されている短単位情報を用い、品詞が「感動詞-フィラー」と付

表1 フィラー語彙素・出現形毎の各年齢区分の調整頻度 (PMW)

語彙素	出現形	全体	1歳	2歳	3歳	4歳	5歳	6歳	7歳	20歳～
■ あの		3537	0	386	2266	6039	5809	3353	4733	10759
	あの	1545	0	245	710	1242	3274	1563	3641	10402
	あのさ	777	0	0	474	1266	1651	846	546	73
	あのですね	5	0	0	0	24	0	0	0	0
	あのな	10	0	35	34	0	0	0	0	2
	あのね	1200	0	105	1049	3507	883	944	546	282
■ あー	あー	34	0	0	68	49	28	33	46	431
■ うんと		250	0	421	541	268	228	65	137	224
	うんと	226	0	421	406	268	199	65	137	167
	うんとさ	5	0	0	34	0	0	0	0	3
	うんとね	19	0	0	101	0	28	0	0	55
■ うー	うー	5	0	0	0	0	28	0	0	138
■ えー	えー	168	0	0	68	73	427	33	637	1061
■ えーと		2913	143	1717	2436	2411	4983	3972	4050	2060
	えーと	2222	143	1262	1827	1583	4043	2897	3458	1870
	えーとさ	235	0	35	68	463	541	260	0	3
	えーとですね	10	0	0	0	0	28	33	0	0
	えーとね	446	0	421	541	365	370	781	592	189
■ その	その	29	0	0	0	0	0	33	228	2108
■ んー		173	0	175	203	170	142	65	501	891
	んー	173	0	175	203	170	142	65	501	848
	んーね	0	0	0	0	0	0	0	0	43

CEJC-Child 全体と 20 歳以上のいずれも PMW が 100 未満の「あーと」「いー」「おー」「と」を除く

与されているものを対象として集計した<sup>1)</sup>。

## 2.3 フィラーの出現位置

フィラーは、発話の逐次的構築を支えるための時間確保や発話継続の表示といった基本機能を担うが、出現位置によっては、発話順番の取得や聞き手の注意獲得などの相互行為的機能が前景化することもある。そこで、フィラーの発話単位中の出現位置に基づき、フィラーを発話単位の「冒頭」か「非冒頭」かに分類した。このうち「冒頭」には、単位先頭のフィラーだけでなく、それに直接後続するフィラー群も含めた。以下に例を示す。なお、発話単位は転記テキスト策定時に [10] に準拠して認定した。

【冒頭の例】会話 ID: Y009\_046, 話者 ID: Y009\_003 (3;1)  
えっとねー うーんとねー ピンクになるよ。

【非冒頭の例】会話 ID: Y001\_086d, 話者 ID: Y001\_000 (4;11)  
これ あのね 夢だったみたいー。

1) CEJC や CEJC-Child では、フィラーは転記テキストの段階で連体詞や感動詞（感動詞一般）と区別しており、品詞情報に反映されている。ただし実際の会話データでは、フィラーと感動詞との区別が困難な事例も存在する。とくに感嘆やためらいなどを表す「あー」「えー」「んー」などの形式が発話の冒頭に出現する場合、文脈や音調によってはフィラーと感動詞の両方の機能を併せ持つ場合があり、一意に特定できないものも多い。このような場合、感動詞としての解釈を優先した。その結果、本研究の分析対象にはフィラー的機能を有する「あー」「えー」「んー」が一部含まれていない可能性がある。また、「なんか」など談話的にはフィラー的な機能を果たす場合がある表現であっても、上記形式に該当しないものは、本研究では対象外となっている。

## 3 形式と出現位置の年齢別分析

分析対象とするフィラーは、基本要素の語彙素（辞書の見出しに相当する「あの」「えーと」など）と、派生要素を含む出現形（「あのさ」「あのですね」など）を単位に集計し、各年齢区分の 100 万語あたりの調整頻度 (PMW) を算出した (表 1)。

まず、フィラー形式および年齢区分の関係について概観する。子ども全体において出現頻度が高かったのは、語彙素が「あの」である形式（以下「あの系」、他の語彙素についても同様に表記）、およびえーと系であった。

あの系は、[5] が分析対象としたあの系・えーと系・んー系の 3 形式の中で、5 名の子ども (2;0-4;11) すべてにおいて最も頻繁に観察された形式であり、本研究でも同様に最も高い出現頻度を示した。あの系は 2 歳以降に出現しており、出現開始時期も [5] と一致する。出現形別に見ると、出現頻度と月齢との関係には差異が認められる。20 歳以上の成人も含めて大局的に見ると、単独の「あの」は細かな増減を伴いながらも、子どもから大人にかけて出現頻度が増加する傾向を示す。一方、子どもが注意獲得などに用いる形式と想定される「あのね」や「あのさ」は、それぞれ 5 歳、6 歳前後に出現頻度のピークを迎え、以降は低下し、成人ではあまり用いられない傾向が見られる。

えーと系は、本研究で扱った形式の中で唯一、2 歳未満 (1;11) の段階で観察されたものであり、子

表2 年齢層別・出現位置別にみたフィラー語彙素の調整頻度 (PMW)

語彙素	位置	全体	1歳	2歳	3歳	4歳	5歳	6歳	7歳	20歳～
あの	冒頭	2217	0	351	1894	4651	2449	2149	2412	3610
あの	非冒頭	1320	0	35	372	1388	3360	1205	2321	7148
あー	冒頭	19	0	0	34	49	28	0	0	68
あー	非冒頭	14	0	0	34	0	0	33	46	364
うんと	冒頭	197	0	351	372	170	228	65	137	135
うんと	非冒頭	53	0	70	169	97	0	0	0	89
うー	冒頭	0	0	0	0	0	0	0	0	49
うー	非冒頭	5	0	0	0	0	28	0	0	90
えー	冒頭	77	0	0	34	0	171	33	364	305
えー	非冒頭	91	0	0	34	73	256	0	273	761
えーと	冒頭	2313	143	1759	1899	3787	3451	2958	920	990
えーと	非冒頭	600	0	677	511	1196	521	1092	0	1072
その	冒頭	5	0	0	0	0	0	0	46	402
その	非冒頭	24	0	0	0	0	0	33	182	1700
んー	冒頭	72	0	70	34	146	57	33	137	358
んー	非冒頭	149	0	140	169	73	85	65	637	533

ども全体における出現頻度もあの系に次いで高かった。えーと系の使用開始時期について、[3]は小学生期を想定しているが、[5]はより早期(2;0-2;5)での出現を指摘しており、本研究の結果もこれを支持するものである。出現形別に、成人も含めて大局的に傾向を見ると、「えーと」「えーとさ」は5歳頃、「えーとね」は6歳頃をピークに、その後、減少する傾向が見られる。[11]は、『日本語話し言葉コーパス』における講演データを対象とした分析で、えーと系の使用率が低いほど「流暢な」「上手な」「話し慣れた」印象を受けることを指摘している。えーと系の使用が発話のつたなさに関わるとするならば、子どもから成人にかけてその使用が減少するという本研究の結果は、この指摘と整合的である。

んー系については、[3]により、対象とした月齢帯で最も早期(3;0-5;0)に独占的に観察された形式(「んー」「んーと」「とー」)として報告されている。しかし本研究では、4歳未満の各年齢区分で最も用いられたのは「えーと」であり、[3]とは異なる結果となった。[3]が指摘するように、んー系は発音しやすい形式である一方で、発話中の息の漏れや詰まりによって頻繁に生じる音でもあり、フィラーとの区別が困難な事例が多く見られた。本研究ではこれらをフィラーと認定せず分析対象外としたため、この点が結果の差異に影響している可能性がある。

最後に、あー系、うー系、えー系など母音を引き伸ばす形式のフィラーおよびその系についてみる。これらのフィラーは子ども全体では出現頻度は低く、成人で高くなる傾向を示している。表に示していないが、いー系やおー系も同様の傾向が確認された。[1]は「えー」が講演場面で多く用いられる形式

であることを指摘しており、[12]も、講演において「いー」「うー」「おー」が20代から80歳以上まで年齢とともに徐々に増加すると報告している。これらの先行研究を踏まえると、母音引き伸ばし型のフィラーは、成長とともに使用されるようになる形式である可能性がある。

続いて、フィラーの発話中における出現位置の特徴について検討する(表2)。[5]が扱ったデータ(2;0-4;11)では、あの系、えーと系のいずれにおいても冒頭での使用が優勢であることが指摘されており、本研究における同年齢帯の結果もこれと一致している。年齢範囲を5歳以降まで拡張して見ると、あの系では、2歳では冒頭使用が圧倒的に多いものの、年齢の上昇に伴い非冒頭での使用が徐々に増加し、5歳以降では冒頭と非冒頭が拮抗する傾向が見られる。さらに成人では非冒頭での出現が冒頭を大きく上回るようになる。これに対しえーと系については、2~7歳の範囲では一貫して冒頭使用が優勢である一方、成人では非冒頭での使用も増加し、両者が拮抗する傾向を示す。このような冒頭使用から非冒頭使用への移行、あるいは非冒頭使用の増加傾向は、あー系においても観察される。

以上の結果から、フィラーの出現位置、ひいてはそれに反映されるフィラー使用の役割は、発達に伴って変化していく可能性が示唆される。一方、うー系およびその系については、5~6歳以降に出現が確認されるものの、当初から非冒頭での使用が多く、この傾向は成人においても大きくは変化しない。このようにフィラーの中には、発達に伴って出現位置や役割が大きく変化するものだけでなく、比較的安定した使用傾向を示すものも存在する。

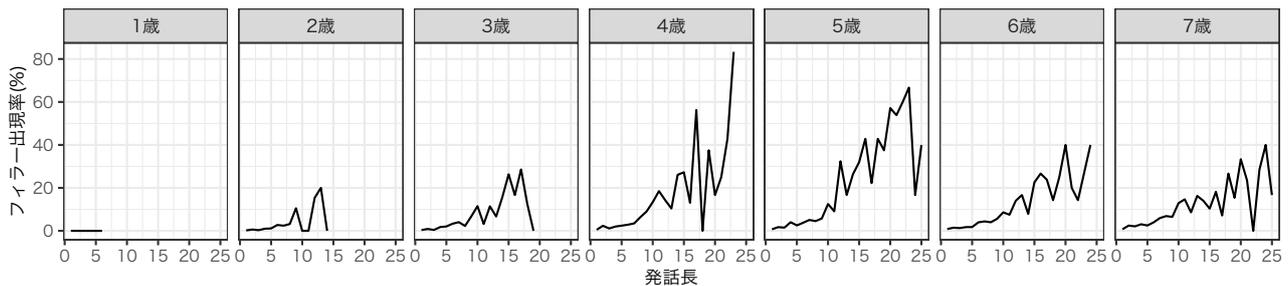


図1 年齢層別に見た発話長とフィラー出現率の関係（各年齢層・発話長の頻度が5以上に限定してプロット）

#### 4 年齢・発話長とフィラー出現率

子どもの認知的成長により発話生成に係るフィラーは減ると考えられる。一方、成長とともに長い発話をするようになり、長い発話では認知的負荷によりフィラーが増えるとも考えられる。よってフィラーの増減は、年齢だけでなく発話長との関係も含めて検討する必要がある。

そこで、年齢ごとに発話長（発話単位に含まれる短単位数）と当該発話にフィラーを含む確率（フィラー出現率）との関係を調べたところ、年齢の上昇に伴いより長い発話を産出できるようになること、また同じ発話長であっても年齢が高くなるにつれフィラー出現率が低下する傾向にあることが確認された（図1）。

そこで、フィラー出現率が年齢と発話長によっていかに変化するかを統計的に検討するため、発話単位にフィラーを含むか否かを二項応答変数とし、一般化線形混合モデルによる分析を行った。固定効果として年齢と発話長を設定した。年齢と発話長は中心化を行った。発達的变化が線形とは限らないことを踏まえ、直交多項式による2次項を用いて非線形効果を検討した。また個人差を考慮するため話者をランダム切片としてモデルに含めた。モデル推定にはlme4パッケージのglmer関数を用いた。

分析の結果、発話長はフィラー出現率に対して正の効果を示した（Estimate = 0.183,  $z = 30.51$ ,  $p < .001$ ）。発話が長くなるほどフィラー出現率が有意に高くなるという、いわば当然の結果である。また話者をランダム切片として含めた結果、話者間にはフィラー出現率に中程度の個人差が認められた。

年齢効果については、直交多項式の一次成分は有意ではなかったが、二次成分は有意であった（ $p < .001$ ）。この結果は、フィラー出現率の発達的变化が単純な線形変化ではなく、非線形的なパターンを示すことを示唆する。固定効果のみを用いて年齢に対

する予測確率曲線を描いたところ、(1) およそ24～60か月にかけてフィラー出現率は緩やかに上昇し、(2) 約60か月以降に急速に低下する傾向が確認された（図2・平均）。さらに(3) この低下傾向は短い発話ではほとんど見られず、長い発話において顕著であった（図2・短い長い）。

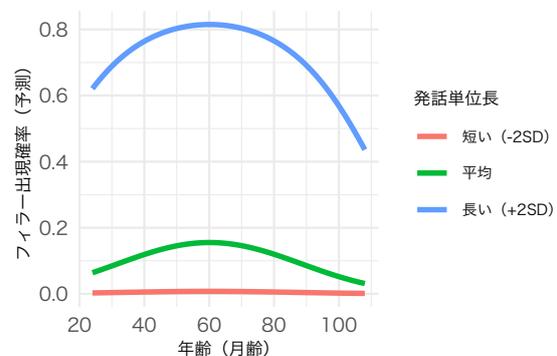


図2 年齢に対するフィラー出現率の予測曲線

以上の結果から、フィラー使用は発達に伴って一様に増減するのではなく、特定の時期を境として変化し、その様相は発話長によって大きく異なることが示された。短い発話では、もともとフィラー出現率が低く、年齢による大きな変化はほとんど見られない。一方、長い発話では、24～60か月にかけてフィラー出現率が緩やかに上昇した後、60か月以降に急速に低下するという発達の推移が確認された。

また話者をランダム切片として含めたモデルにおいて、フィラー出現率には話者間で有意な個人差が認められたことから、フィラー使用は年齢や発話長といった発達の・構造的要因だけでなく、話者固有の発話スタイルやコミュニケーション方略にも影響を受けている可能性が示唆される。

今後は、フィラーの形式や出現位置といった側面も考慮し、年齢や発話長との関係をより統合的に検討する。また子どものフィラー使用の機能的側面や相互行為上の役割をより精緻に分析する。

## 謝辞

本研究は国語研究所共同研究プロジェクト「多世代会話コーパスに基づく話し言葉の総合的研究」および科研費 23K25327,23K17586,22K13109,20H05617 の成果である。

## 参考文献

- [1] 山根智恵. 日本語の談話におけるフィラー. くろしお出版, 2002.
- [2] 定延利之, 田窪行則. 談話における心的操作モニター機構——心的操作標識「ええと」と「あの(ー)」. 言語研究, No. 108, pp. 74–93, 1995. [https://doi.org/10.11435/gengo1939.1995.108\\_74](https://doi.org/10.11435/gengo1939.1995.108_74).
- [3] 塩沢孝子. 日本語の Hesitation に関する一考察. ことばの諸相, pp. 151–166. 文化評論出版, 1979.
- [4] 伝康晴. <こども>のフィラー. ことばと文字, Vol. 14, pp. 34–42, 2021.
- [5] KUBOTA Mihoko. Development of the use of fillers by japanese-acquiring children. 東京造形大学研究報, No. 24, pp. 153–171, 2023. <https://zokei.repo.nii.ac.jp/records/140>.
- [6] 友定賢治. 幼児のことばの非流暢：E 児データによるケーススタディー. 定延利之ほか (編), 流暢性と非流暢性, pp. 471–492. ひつじ書房, 2024.
- [7] 小磯花絵, 石本祐一, 居關友里子, 江口典子, 柏野和佳子, 川端良子, 田中真理子, 田中弥生, 西川賢哉. 子ども版日本語日常会話コーパスモニター版の構築. 言語処理学会第 31 回年次大会発表論文集, pp. 3525–3528, 2025. [https://www.anlp.jp/proceedings/annual\\_meeting/2025/pdf\\_dir/B9-5.pdf](https://www.anlp.jp/proceedings/annual_meeting/2025/pdf_dir/B9-5.pdf).
- [8] 小磯花絵, 天谷晴香, 居關友里子, 白田泰如, 柏野和佳子, 川端良子, 田中弥生, 伝康晴, 西川賢哉, 渡邊友香. 『日本語日常会話コーパス』設計と特徴. 国立国語研究所論集, Vol. 24, pp. 153–168, 2023. <https://doi.org/10.15084/00003692>.
- [9] 白田泰如, 川端良子, 西川賢哉, 石本祐一, 小磯花絵. 『日本語日常会話コーパス』における転記の基準と作成手法. 国立国語研究所論集, No. 15, pp. 177–193, 2018. <https://doi.org/10.15084/00001602>.
- [10] JDRI. 発話単位ラベリングマニュアル version 2.1, 2017. <https://www.jdri.org/resources/manuals/uu-doc-2.1.pdf>.
- [11] 小磯花絵. 「場つなぎ表現」が聞き手に与える印象. 情報通信ジャーナル, Vol. 24, No. 3, p. 28, 2006.
- [12] 小磯花絵. 「『日本語話し言葉コーパス』『日本語日常会話コーパス』利用のべし・べからず」. 2023 年度 NINJAL シンポジウム「コーパス利用のべし・べからず」, 2023.

## A 付録

表3 調査協力世帯・調査対象児の基本情報

世帯 ID	話者 ID	性別	収録開始時の月齢	収録期間	同居家族	備考
Y001	Y001_000 (Y001_017)*	女 (女)	2歳6ヶ月 (0歳0ヶ月)	53ヶ月 (29ヶ月)	父・母	父は CEJC の調査協力者
Y002	Y002_000	女	5歳8ヶ月	32ヶ月	父・母	日韓バイリンガル
Y005	Y005_000 (Y005_007)*	女 (男)	1歳7ヶ月 (0歳2ヶ月)	35ヶ月 (19ヶ月)	父・母	日中バイリンガル
Y006	Y006_000 Y006_004	女 男	1歳6ヶ月 6歳6ヶ月	35ヶ月	父・母 姉(小学生)	
Y008	Y008_000	女	0歳9ヶ月	16ヶ月	父・母	
Y009	Y009_003 Y009_000	女 男	1歳2ヶ月 4歳2ヶ月	24ヶ月	父・母	
Y010	Y010_000 Y010_003	男 男	0歳10ヶ月 3歳8ヶ月	39ヶ月	父・母	
Y011	Y011_001 Y011_000	男 男	1歳6ヶ月 4歳11ヶ月	13ヶ月	父・母	

\* 調査開始後に誕生。本分析では対象外とする。

表4 年齢区分別のデータ規模

年齢区分	異なり話者数	語数
1歳 (1;0 - 1;11)	6	6,977
2歳 (2;0 - 2;11)	7	28,530
3歳 (3;0 - 3;11)	6	29,562
4歳 (4;0 - 4;11)	7	41,065
5歳 (5;0 - 5;11)	5	35,120
6歳 (6;0 - 6;11)	6	30,718
7歳 (7;0 - 7;11)	4	21,975
20歳以上	469	1,516,776