

中国における新型コロナウイルス感染症に関するフェイクニュースの構造

ーテキストマイニング分析を中心にー

Structure of fake news about COVID-19 in China

-Analysis using text mining method-

¹李 瑞鵬/Ruipeng LI・²松野 良一/Ryoichi MATSUNO

¹中央大学大学院総合政策研究科 博士後期課程・²中央大学国際情報学部 教授

[Abstract]

This thesis elucidates the overall structure of fake news related to COVID-19 in China through text mining. The study analyzed 806 pieces of information that the Chinese social media platform WeChat fact-checked and determined as fake news (from January 2020 to December 2022). This paper reveals several aspects regarding the occurrence and processes of fake news during a pandemic: Initially, fake news emerged concerning the novel virus itself. Subsequently, fake news about preventive measures arose. Fake news disguised as "announcements" from public institutions continued to emerge. Fake news primarily spread through videos on the internet, rapidly disseminating in close friend circles. In terms of chronology, fake news first emerged from Wuhan and then spread to Shanghai and other cities. While fake news about affected cities was expected to appear, those on prevention methods related to daily activities—such as drinking and eating—arose when the infection proliferated. Delayed responses may lead to a significant increase in fake news related to announcements from institutions. This suggests that citizens are eager for information from government and local authorities, emphasizing promptness and transparency in the information released by public institutions.

[キーワード]

フェイクニュース、新型コロナウイルス感染症、ファクトチェック、KH Coder、テキストマイニング

1. はじめに

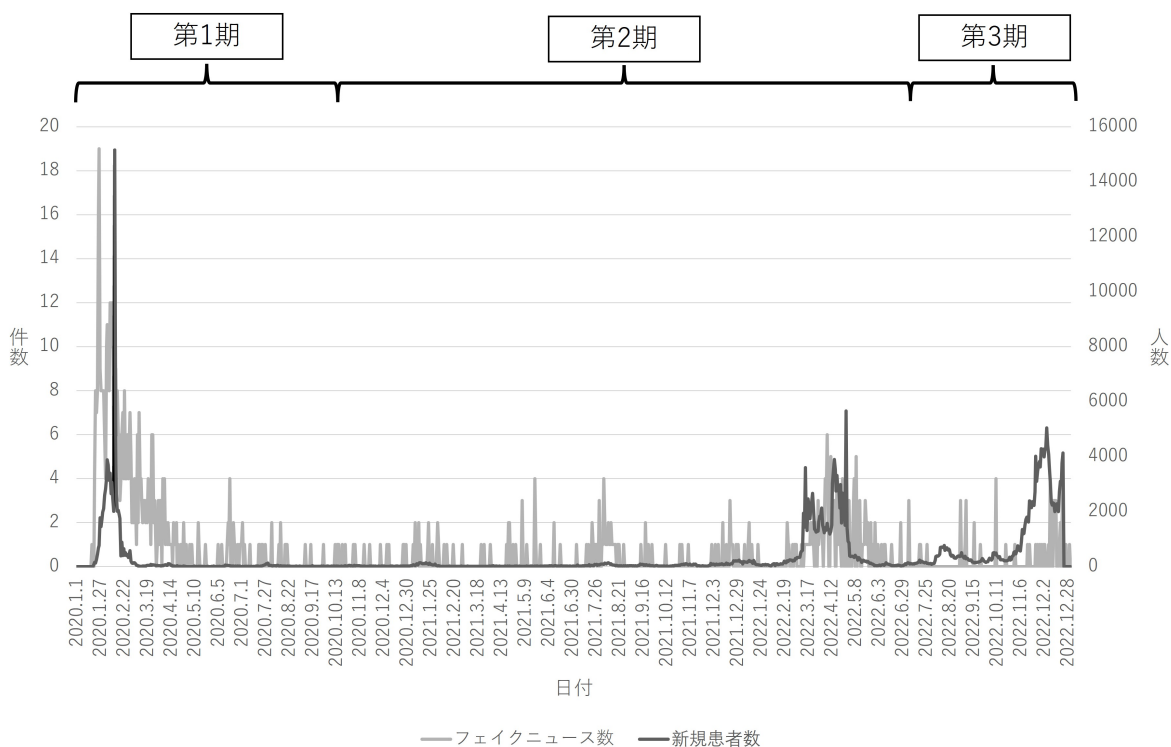
1.1 問題と背景

本論文は、中国で発生した新型コロナウイルス感染症に関するフェイクニュースの全体的構造をテキストマイニング分析によって明らかにすることを目的とした。研究対象とした期間は、中国武漢市で初めて患者が発生した2020年1月から、中国政府（中国衛生健康委員会）が「本日をもって新型コロナウイルスに関する情報の発表を停止する」とした2022年12月までの3年間。つまり、「発生」から「終息」までの3年間に発生したフェイクニュースを分析対象とした。

図-1は、中国衛生健康委員会が発表した新規患者数（黒色実線）と、中国のファクトチェックサイトである微信「較真」がフェイクニュースと断定した情報数（灰色実線）の変化である。「微信」は中国版LINEと呼ばれるソーシャルメディアであり、2020年1月18日から新型コロナウイルス感染症に関する情報を専門にファクトチェックするサイト微信「較真」を開設し、その結果を公開している。

図-1によると、新規患者数が急増している時期が3か所あることがわかる。1番目は2020年1月から2月にかけて、2番目は2022年3月から5月にかけて、3番目は2022年11月から12月にかけて、である。

1番目の山は、武漢市で発生し全土に拡大していった時期、2番目の山は、上海市で大規模な流行が発生した時期、3番目の山は、広東省、北京市など複数の地域で拡大した時期である。



図一 中国政府が発表した新規患者数と微信「較真」がフェイクニュースと判定した情報数（筆者作成）

図一を概観すると、新規患者数とフェイクニュース数は、かなりの部分で連動していることがわかる。

本論文ではまず、2020年1月1日～2022年12月31日までの3年間に於いて、微信「較真」がフェイクニュースと判断した新型コロナウイルス感染症に関する情報を収集し、テキストマイニングツールである KH Coder を使って分析する。

さらに、全期間を、第1期（2020年1月1日～2020年9月30日）、第2期（2020年10月1日～2022年6月30日）、第3期（2022年7月1日～2022年12月31日）に分け、その3期間の差異について明らかにする。

第1期と第2期の境界を2020年9月30日にした理由は、この日までに新規患者数が一度ゼロになり、一旦感染状況が落ち着いたと判断できたからである。そして、第2期は上海市で感染が拡大した時期を含んでおり、新規患者数は5659人まで上昇した。しかし、2022年6月下旬には、10～30人で推移し、感染拡大は一旦落ち着いたと判断した。このため、6月30日までの第2期と設定した。また、中国政府（中国衛生健康委員会）は、2022年12月25日をもって、新型コロナウイルス感染症に関する情報の発表を停止した。このため、第3期を2022年12月31日までと設定した。

1.2 先行研究

ここで、本論文と関連する先行研究について、レビューしておきたい。

アメリカの心理学者 G.W.オルポートとポストマンは『デマの心理学』（1952）で法則「 $R \sim I \times A$ 」を提唱した。Rはrumorのことで、デマ（うわさ）の流布量である。Iはimportanceのことで、内容の重要性である。Aはambiguityのことで、内容の曖昧さである。つまり、ある事案について、内容の重要性と曖昧さが高ければ高いほど、うわさが発生し拡散しやすいとした。

中森（2020）は、災害時の流言の伝播について、旧来のパターンである直接人から人へ伝えられるという形から、インターネット・SNSを介して人から人へ伝えられる形になったと指摘した。

福長（2020）は、新型コロナウイルス感染症発生と同時に発生したトイレトペーパー買いだめ事件について、フェイクニュースの流布過程を分析した。まず、感染症が発生した後、マスク不足の状況が出現した。ユーザーの中には、マスク不足の状況をオイルショック時のトイレトペーパー買いだめ騒ぎと関連させて SNS に投稿した者もいた。これがオイルショック時の騒ぎを想起させ、人々は強い不安を抱き、トイレトペーパー買いだめ行動につながったと考察した。

銭 (2020) は、微信「較真」がファクトチェックを行った、2020年1月18日から2020年3月26日までの528件の情報について、定性的コーディング分析を使って分類と分析を行った。その結果、新型コロナウイルス感染症が深刻なエリアほど、フェイクニュースの数も増加したと報告している。銭は感染症が厳しい状況になればなるほど、人々の不安が増しフェイクニュースの拡散を加速させると考察している。しかし、銭の研究は、新型コロナウイルス感染症の初期のデータのみを基にした分析であり、全体像を明らかにするまでには至っていない。

李・松野 (2022) は、2020年1月1日から2020年9月30日の期間における中国における新型コロナウイルス感染症に関する情報について、微信「較真」がフェイクニュースと判定した489件の情報を、KJ法で分類した。その結果、新型コロナウイルス感染症に関するフェイクニュースは、「医療公衆衛生系」と「社会行動系」の2つの大グループに分類されることなどを明らかにした。

さらに李・松野 (2023) は、2020年1月1日～2022年6月30日までの2年半において、微信「較真」がフェイクニュースと判断した新型コロナウイルス感染症に関する情報751個を収集し、KJ法を使って構造を明らかにした。全期間を、第1期 (2020年1月1日～2020年9月30日) と第2期 (2020年10月1日～2022年6月30日) に分け、その2期間の差異について明らかにした。第1期は「感染予防と治療」に関するフェイクニュースが多く発生し、第2期では「社会ニュース」に関するフェイクニュースが多く発生したこと。また、第2期には「詐欺の疑い」と思われるフェイクニュースが登場したことなどを明らかにした。しかし、この研究は定性的分析方法であるKJ法を使用しており、期間も2年半であった。

本論文の独自性・新規性は、①新型コロナウイルス感染症の発生から終息までの全期間、つまり2020年1月1日から2022年12月31日までの3年間を研究対象としたこと、②テキストマイニングを使用して中国における新型コロナウイルス感染症に関するフェイクニュースの全体構造を明らかにしたこと、③3年間を前期 (第1期)、中期 (第2期) と後期 (第3期) に分け、テキストマイニングを使って時系列的变化、および3期間ごとの特徴と相互関係性を明らかにしたこと、の3点である。

2. 本論文の分析対象と目的

2.1 微信「較真」がファクトチェックした情報を研究対象とした理由

本研究は、中国のSNSの1つである微信が運営するファクトチェックサイト、微信「較真」が公開したフェイクニュースを分析対象とした。その理由について、最初に説明しておきたい。

1つ目は、微信「較真」がファクトチェックの対象とする情報の収集と選別の方法が明確であること。

中国のITメディア企業「テンセント社」が運営する「微信」は、新型コロナウイルス感染症に関する情報をファクトチェックする専門サイト「較真」を設置し、2020年1月18日から、その結果を公開している (図-2)。

ファクトチェックの手順は、以下の通りである。微信「較真」は、中国の主なSNS (微博²、微信、知乎³など) で流れている真偽不明な情報を収集し、フェイクの可能性のあるものをスクリーニングする。そして、ピックアップした情報について、問題の重要性、潜在的な危害、緊急性、曖昧さ、のレベルを基準に、さらに選別する。その後、選別された情報を注目度の高い順に並べて最終的なチェック対象を定める。選別された情報がフェイクかどうかの判定については、微信「較真」と協力関係にある「科技日報」「騰訊新聞 (テンセントニュース)」「中国医師協会健康伝播工作委员会」など約100の団体、さらに、医療衛生、食品安全、科学技術、国際問題などの領域の専門家に依頼し、助言を受けて判断を下す。

この方法であれば、ユーザーの注目度が高い情報から順番でファクトチェックを行い、その結果を効率よく公表することが可能であり、それによってフェイクニュースの拡散に素早く対処できるとしている。



図-2 COVID-19 関連の情報に特化してファクトチェックを行っている微信「較真」

2つ目に、中国のユーザーは、新型コロナウイルス感染症に関する情報を、微信から最も多く得ているという調査結果があること。

中国国家情報センターと南京大学ネットワーク伝播研究院は、2020年1月20日から2月13日の間に、「2020“新型コロナウイルス肺炎”公衆認知と情報伝播調査研究」を行った。調査方法は、主に定量分析のアンケート調査と定性分析のインタビュー調査である。そのうち、アンケート調査は、中国大陸31個の省、市や自治区にいる3202人を対象として質問票を配布し、有効回答率は93.7%だった。結果は、新型コロナウイルス感染症に関する情報を取得する際の入手ルートは、微信が62%であり最多であった(図-3)。

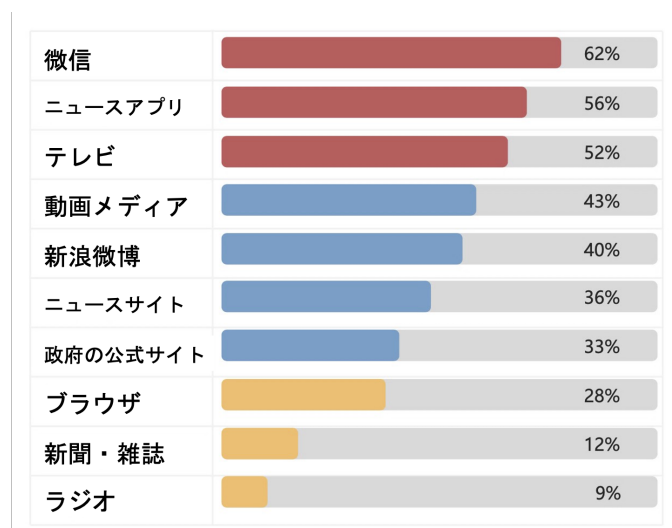


図-3 新型コロナウイルス感染症の関連情報を取得するルート

3つ目は、微信「較真」は、「騰訊新聞」および「中国医師協会健康伝播工作委员会」と協力してファクトチェックするため、信頼度が高いこと。

「2020“新型コロナウイルス肺炎”公衆認知と情報伝播調査研究」の調査によれば、「騰訊新聞」は情報源として、2番目に信頼できるとされた(1番目は人民日報)。また、「中国医師協会健康伝播工作委员会」は「中国医師協会」に属する機関で、同協会は、中国国務院(日本では内閣)に属する行政部門であるため、一定以上の信頼性があると考えられている。

以上の3つ理由から、微信「較真」がファクトチェックし結果を公表した情報を本論文の研究対象とした。

2.2 研究目的

本論文の目的は3つある。

1つ目は、微信「較真」によってフェイクニュースと判定された3年間の情報を、テキストマイニングツールKH Coderを使って、全体的な構造や特徴を明らかにすること。

2つ目は、KH Coderによる外部変数を使った分析で、3期間のフェイクニュースの時系列的変化を明らかにすること。

3つ目は、KH Coderの対応分析を使用して、3期間ごとのフェイクニュースの特徴を明らかにすること。

3. 微信「較真」がファクトチェックした情報の事前集計

3.1 事前集計の結果

微信「較真」が、新型コロナウイルス感染症に関する情報についてファクトチェックした結果を収集し集計した。期間は、2020年1月1日から2022年12月31日までの3年間。

微信「較真」がファクトチェックした情報は、1027個。判定結果は、「真実」が75個、「判断できない」が146個、「フェイクニュース」は806個であった。本研究では、この「フェイクニュース」と判断された806個の情報を研究対象とした。

表-1 微信「較真」による判定結果（3年間、筆者作成）

判定結果	数
真実	75
判断できない	146
フェイクニュース	806
合計	1027

3.2 期間ごとの事前集計の結果

次に、806個のフェイクニュースを3期に分けた。第1期489個、第2期262個、第3期55個となった。

表-2 微信「較真」がファクトチェックした全期間を3期に分けた集計結果（筆者作成）

	期間	フェイクニュース数
第1期	2020年1月1日～2020年9月30日	489
第2期	2020年10月1日～2022年6月30日	262
第3期	2022年7月1日～2022年12月31日	55

4. 中国における新型コロナウイルス感染症に関するフェイクニュースの全体像

4.1 方法と手続き

微信「較真」がフェイクニュースと判断した806個の情報の内容を研究対象とした。テキストマイニング分析には、KH CoderのVersion3.Beta.07dを使用した。

はじめに、この806個のフェイクニュースの内容を確認し、絵文字などのパソコンで読み込み不可能な表記を削除し、誤字脱字の修正を行うことでデータのスクリーニングを行った。その後、この806個のフェイクニュースをExcelに取り込み、ファクトチェック順に並べた。さらに、「新型コロナウイルス」「コロナウイルス」などの表記揺れや同義語を修正した。分析用データとして整ったところで、KH Coderに読み込ませ、前処理を行った。その後、以下の手順で分析を行った（表-3）。

表-3 具体的な操作手順と変数操作の設定

分析内容	操作手順と変数操作の設定
① 頻出語リストの作成（4.2参照）	「ツール」「抽出語」「抽出リスト」を選択する
② 全期間の共起ネットワーク図作成（4.3参照）	「ツール」「抽出語」「共起ネットワーク」を選択し、「最小出現数」を20に変更、「強い共起関係ほど濃い線に」をクリックする
③ 「期間」を外部変数とした共起ネットワーク図作成（4.4参照）	「ツール」「抽出語」「共起ネットワーク」を選択し、「最小出現数」を20に変更、「共起関係（edge）の種類」のところに「語-外部変数・見出し」「期間」を選ぶ。そして「強い共起関係ほど濃い線に」をクリックする
④ 「期間」を外部変数とした対応分析（4.5参照）	「ツール」「抽出語」「抽出リスト」を選択し、「最小出現数」を20に変更、「抽出語×外部変数」のところに「期間」を選択する

4.2 頻出語リスト作成に関する分析結果と考察

まず、微信「較真」がフェイクニュースと判定した情報806個について、KH Coderを使って頻出語リストを作成した。

使用語彙の解析の結果、段落数806、文数1168が確認された。また、総抽出語数（分析対象ファイルに含まれているすべての語の延べ数）は30695、異なり語数（何種類の語が含まれていたかを示す数）は6644であった。その中で、分析に使用される語（助詞や助動詞等のような文章にでも現れる一般的な語が除外された数）として15501語、異なり語数5460が抽出された。

表-4は、全期間のフェイクニュースの頻出語リストである。最小出現数が25回以上の上位66語とその出現回数を示している。

表-4 微信「較真」がフェイクニュースと判断した情報の単語上位 66 語（最小出現数が 25 回以上）

順位	抽出語	出現回数	順位	抽出語	出現回数
1	新型冠状病毒（新型コロナウイルス）	336	34	开学（学校が始まる）	36
2	病毒（ウイルス）	221	35	是真的吗（本当か）	36
3	人（人）	170	36	导致（引き起こす）	35
4	流传（流布）	147	37	核酸检测（PCR 検査）	35
5	网上（ネット上）	121	38	美国（米国）	35
6	视频（動画）	114	39	阳性（陽性）	35
7	肺炎（肺炎）	112	40	可（できる）	34
8	新型（新型）	107	41	网络（インターネット）	34
9	感染（感染）	106	42	发布（公的機関の発表）	33
10	疫情（感染状況）	91	43	通知（告知）	33
11	网传（ネット伝播）	82	44	隔离（隔離）	33
12	医院（病院）	79	45	预防（予防）	33
13	能（できる）	76	46	引发（引き起こす）	32
14	上海（上海）	72	47	不要（～してはいけない）	31
15	消息（メッセージ）	67	48	新冠疫苗（コロナワクチン）	31
16	武汉（武漢）	66	49	物资（物資）	31
17	疫苗（ワクチン）	55	50	病人（病人）	31
18	出现（出現）	50	51	需要（需要）	31
19	口罩（マスク）	50	52	不能（できない）	30
20	说（言う）	49	53	传言称（うわさによると）	29
21	网友（ネットユーザー）	46	54	北京（北京）	29
22	发现（発見）	44	55	图片（画像）	29
23	消息称（情報によると）	43	56	居民（居住者）	29
24	人员（関係者）	42	57	开始（開始）	29
25	传播（伝播）	42	58	确诊（診断）	29
26	钟南山（鍾南山）	41	59	人体（人体）	28
27	显示（示す）	40	60	内容（内容）	28
28	进行（行う）	40	61	风险（リスク）	28
29	喝（飲む）	39	62	信息（情報）	27
30	小区（団地）	39	63	照片（写真）	27
31	吃（食べる）	38	64	中国（中国）	26
32	截图（スクリーンショット）	38	65	出门（外出）	25
33	患者（患者）	37	66	朋友圈（モーメンツ）	25

この頻出語リストを見ると、最も出現回数が多かったのは、「新型コロナウイルス」そのものに関する語であった。「新型コロナウイルス (336 回)」「ウイルス (221 回)」「肺炎 (112 回)」「新型 (107 回)」「感染 (106 回)」などである。地球上に突然出現した未知のウイルスについて、多くの人が知りたいと思い、それがフェイクニュースを生み出した可能性が大きい。

フェイクニュースの情報源・伝達経路に関する語としては、「ネット上 (121 回)」「ネット伝播 (82 回)」「言う (49 回)」「インターネット (34 回)」「告知 (33 回)」「モーメンツ (25 回)」が頻出した。インターネットやロコミが、フェイクニュースの主要な拡散ルートであったことがわかる。

特徴的だったのは、政府による公式発表である「告知」が 33 回も出現した。ロックダウンされ情報が乏しく、

市民が公的機関からの「告知」に飢えていたため、逆に「告知」に関するフェイクニュースが多数発生した可能性がある。

最後の「モーメンツ」は、ソーシャルメディア「WeChat」の機能のひとつで、友達で構成されるグループのこと。投稿された写真やメッセージを見ることができるのは友達グループだけであるため、ユーザーは通常よりも信じやすい状態に置かれ、フェイクニュースの拡散速度も上がると考えられる。

フェイクニュースの拡散メディアとして、「動画 (114 回)」「スクリーンショット (38 回)」「画像 (29 回)」「写真 (27 回)」も特徴的である。フェイクニュースは主に動画、スクリーンショット、写真といったメディアを通じて拡散されたことがわかる。特徴的だったのは、「動画」の数が他のメディアを大きく上回ったこと。これは、ロックダウン期間中は外出できないため、人々は「動画」を視聴することが主要な情報収集の手段となっていた。この「動画」の中にフェイクニュースが多数含まれていたと考えられる。

場所や地域に関する頻出語として、「病院 (79 回)」「上海 (72 回)」「武漢 (66 回)」「団地 (39 回)」「北京 (29 回)」が挙げられる。「病院」は治療や隔離に関連しており、「団地」は封鎖に関連している。地域に関する頻出語の中では、「上海」が最も多く、これは上海におけるロックダウン期間中の厳格な封鎖措置が原因であると思われる。

予防法に関する頻出語として、「ワクチン (55 回)」「マスク (50 回)」「飲む (39 回)」「コロナワクチン (31 回)」が挙げられる。中でも、「飲む」という行為を通じて新型コロナウイルスを予防するというフェイクニュースが多く発生した。事例として、「最新の研究によれば、お茶を飲むことは新型コロナウイルス感染症の予防に非常に効果的である」というものがあった。

以上のことから、フェイクニュースに含まれる頻出語は、5つの大きなグループに分けられることがわかった。①「新型コロナウイルス」そのものに関連するグループ、②情報源・伝達経路に関連するグループ、③拡散メディアに関連するグループ、④場所や地域に関連するグループ、⑤予防法に関連するグループ、である。

4.3 全期間の共起ネットワーク分析結果と考察

微信「較真」がフェイクニュースと判断した情報 806 個 (全期間) について、KH Coder を使って共起ネットワーク図の作成を行った。その結果が、図-4 である。

右側のサブグラフは、「新型コロナウイルス・予防法」に関するグループで、「病毒 (ウイルス)」「人」「肺炎」「感染」「新型」「預防 (予防)」「喝 (飲む)」「水」などの語群で構成されている。

武漢市での感染爆発でパニックに陥った人たちは、まず、この未知のウイルス自体について強く知りたいと思ったため、関連するフェイクニュースが発生したと考えられる。事例として、「新型コロナウイルスは SARS ウイルスであり、専門家が記者会見でそう言った」というものがあった。

また、「予防法」に関する事例として、「60 度のお湯をよく飲むと、新型コロナウイルスを予防できる」というものがあった。

さらに上部に破線でつながっているのは、「吃 (食べる)」「阳性 (陽性)」「不能 (できない)」という語のグループである。事例として、「野菜は食べてはいけない。野菜を食べると新型コロナウイルスに感染する」というものがあった。

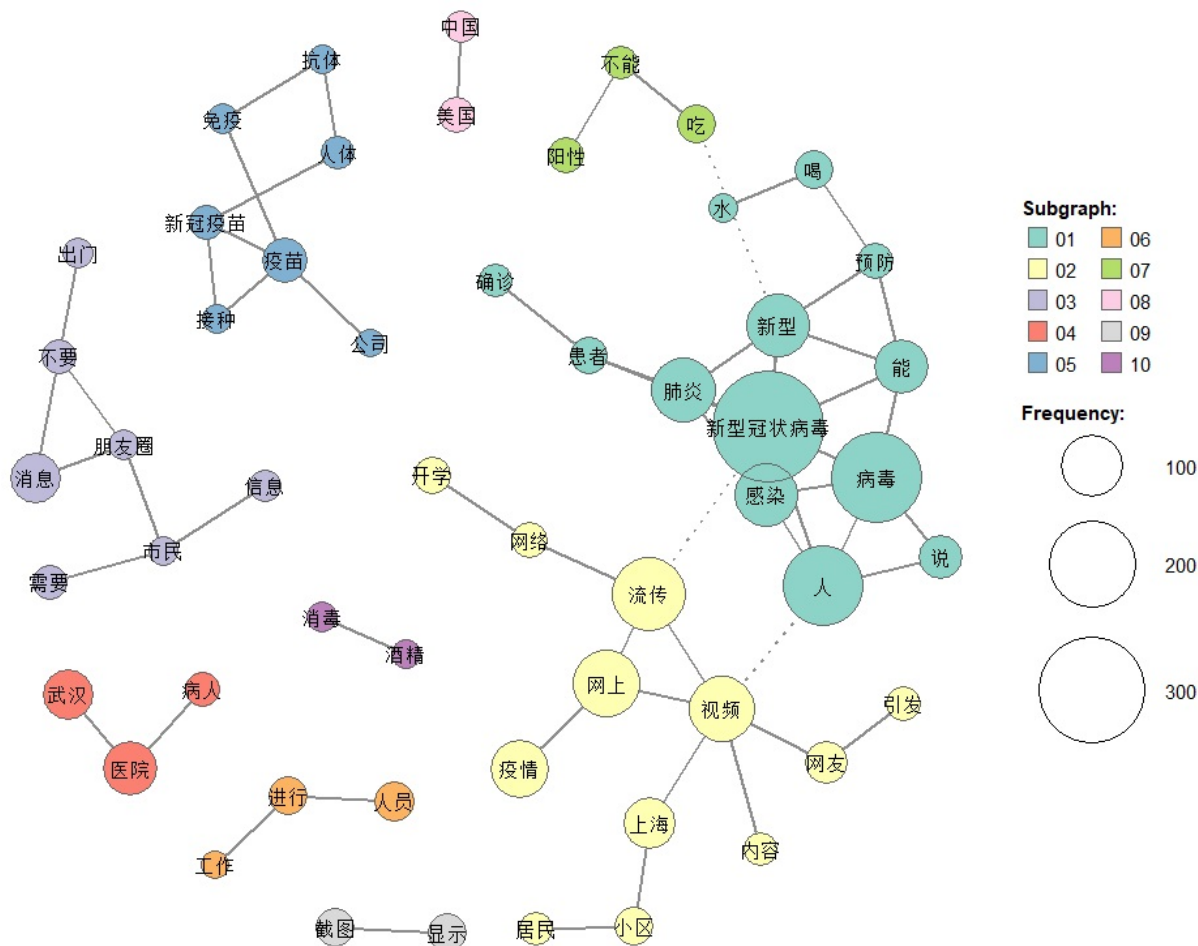


図-4 全期間におけるフェイクニュースと判断された情報806個の共起ネットワーク

右下のサブグラフの中心を構成するのは「流传（流布）」「网上（ネット上）」「视频（動画）」などの語であり、これによると、フェイクニュースは、ネット上の動画によって流布されたことがわかる。また、「视频（動画）」「上海」「小区（団地）」「居民（住民）」が繋がっているが、これはロックダウンされた上海の団地の中で、住民が動画を視聴したため、フェイクニュースが一気に流布したことが推察される。

以上の3つのサブグラフ、「新型コロナウイルス・予防法」グループ、「食べる・陽性・できない」グループ、「流布・ネット・動画」グループ、は破線でつながっており、相互に関連していることがわかる。

次に、独立したサブグラフに注目する。

左側のサブグラフは「朋友圈（モーメンツ、友達グループの意味）」「信息（情報）」「市民」「需要」「出门（出かける）」「不要（～してはいけない）」「消息（メッセージ）」が共起する語群である。中心の「モーメンツ（友達グループ）」を通じて、フェイクニュースが多く流布したことがわかる。また、「出かける」「～してはいけない」が繋がった部分は、政府など公的機関による「告知」に関連するフェイクニュースが含まれていると思われる。事例として、「明日の朝4時に飛行機が消毒液を散布するので、みなさん外出しないでください。お互いに伝え合ってください」というものがあった。

左上のサブグラフは、コロナワクチンとワクチン接種に関する語群である。構成する語は「疫苗（ワクチン）」「新冠疫苗（コロナワクチン）」「接种（接種）」「免疫」「人体」「抗体」「公司（会社）」である。特に、中心の「ワクチン」「コロナウイルス」「接種」の共起関係が強いことが分かる。人々は、ワクチンが人体へ及ぼす副作用、効果について、疑問や不安を持っていた。このことがワクチン接種に関するフェイクニュースを生み出した可能性がある。事例として、「スペインの老人ホームで、ワクチン接種後に全員の陽性反応を確認したと報道された。多くの人々が、新型コロナウイルスの感染はワクチン接種が原因であると考えている」というものがあった。

真上のサブグラフは、「中国」と「美国（米国）」で構成されている。事例として、以下のようなものがあった。

「インターネット上で映像が流れている。中国に向かう米国の医療チームが出発する前に祈っているものであ

る」「1分前に受け取った情報：トランプが突然気を失った！速報！米国大統領ドナルド・トランプ氏が新型コロナウイルスに感染し、病院で治療を受けている。大統領の職務は副大統領のペンスが引き継いでいる。米国全土がパニックに陥っている。中国入国ビザを持つ米国人が中国行きフライトを争っており、航空券の価格は通常の10倍である」「米国では疫病のために多くの人々が死亡し、死体が太平洋に浮かんでおり、海水が汚染され、中国では食塩を生産することができなくなった。みんな急いで市場に行き塩を買いだめしてください。」

中国と米国の関係性がどのように、これらのフェイクニュースに反映されているのかは、これだけでは考察できない。しかし、「中国」と「米国」が対になって共起している点が、中国における新型コロナウイルス感染症に関するフェイクニュースの特徴の1つであることがわかった。

4.4 「期間」を外部変数とした共起ネットワーク分析結果と考察

3期間の特徴と時系列的变化を捉えるために、「期間」を外部変数として共起ネットワーク図の作成を試みた。その結果を、図-5に示した。

まず、第1期、第2期、第3期に共通している語は、中心にある部分である。「新型冠状病毒（新型コロナウイルス）」「流传（流布）」「医院（病院）」「网上（ネット上）」「人」である。これにより、フェイクニュースはネット上で拡散されたことがわかる。ロックダウンによって外出が制限され、人々は主にネットを通じて情報を得る時間が長くなった。このため、フェイクニュースは、ネット上で発生と拡散を繰り返したと考えられる。

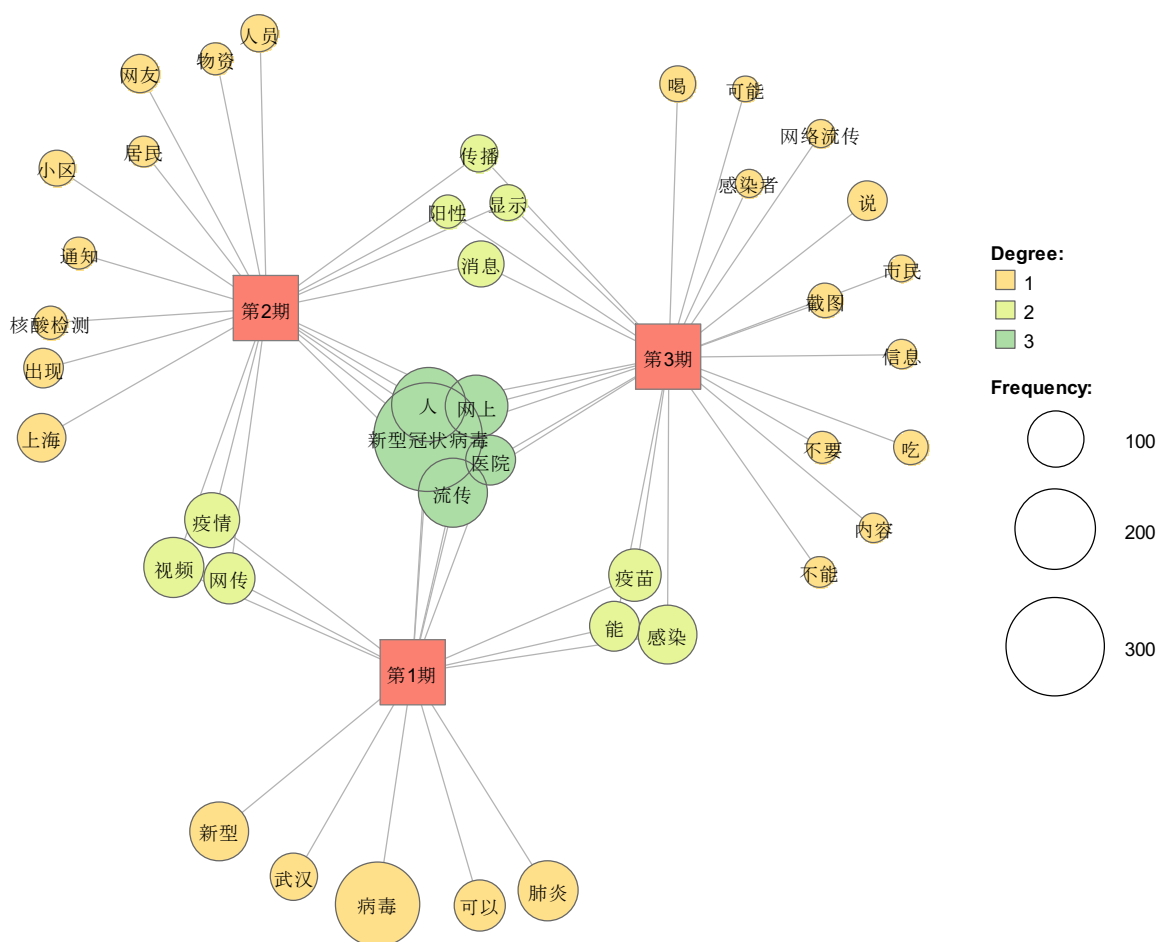


図-5 外部変数「期間」を導入した全期間の共起ネットワーク図

次に、各期別に分析・考察する。

第1期では下側の「新型」「武漢」「病毒（ウイルス）」「肺炎」等の単語が特徴的であった。フェイクニュースの事例を見ると、「武漢の未知の原因の肺炎はSARS ウイルスが引き起こす」などがあった。

「可以」というのは日本語では可能、～ができるという意味をもっている。事例として、「最近、武漢の新型コロナウイルス感染症が拡大している中で、抗インフルエンザ薬のオセルタミビルで感染症を予防できるという話

が出ている」というフェイクニュースがあった。つまり、第1期は、ウイルスそのもの、および予防法に関するフェイクニュースが多かった。

第2期では、上海市で感染が拡大し、都市がロックダウンされた。頻出語は「上海」「小区（団地）」「核酸検査（PCR検査）」「通知（告知）」「物資」などである。「団地」の封鎖や定期的な「PCR検査」が常態化された。

フェイクニュースの事例としては、「3月26日、ネット上で『上海のある老人が団地の封鎖により薬を取りに行けず、飛び降り自殺した』という情報が流れ、その現場の写真が添付されていた」「PCR検査の綿棒には発がん性の物質（蛍光剤など）が含まれており、PCR検査を多く受けると、軽度なら咽頭炎を引き起こし、重度なら他の悪い病気を引き起こす可能性がある」というものがあった。

また、政府の「通知（告知）」の語も頻出語であった。事例として、「重要な告知：明日から警察や紀律検査監査委員会の人々が町をパトロールし、身分証なしで外出している住民を発見した場合、施設に連れて行き、14日間隔離する。1日あたり470元を自己負担する必要がある」というものがあった。

第3期では、中国の複数都市で無症状の感染者や再感染者が増加したため、第1期と同じように予防法に関するフェイクニュースが再び増加した。「喝（飲む）」「可能（できる）」「吃（食べる）」「不要（～してはいけない）」などが頻出語であった。事例として、「専門家が、新型コロナウイルスの感染者に接触した者に対し、症状を軽減するために、通常の半分の量の風邪薬を発症前に摂取すると良いと提案している」というものがあった。

次に各期間の相互関係について考察する。

第1期と第2期では、「视频（動画）」「疫情（感染状況）」「网传（ネット伝播）」という語が共通していた。これにより、第1期と第2期においては、フェイクニュースが動画の形でネット上に拡散され、内容は「感染状況」に関するものであった。事例として、「ネット上で流れている感染症の対策として、深圳（セン）の49の病院で無料のマスクを受け取ることができる」というものがあった。

第2期と第3期においては、「消息（メッセージ）」「显示（示す）」「阳性（陽性）」「传播（伝播）」という語が共通していた。この時期は、団地のロックダウンが行われており、陽性患者の発生と封鎖に関連するフェイクニュースが多く出現している。事例では、「6月8日の夜、『全市の各区や町で、段階的な団地の封鎖措置が実施され、48時間以内に1～2回のPCR検査が行われる。陽性患者が確認された団地は、2+14日間の隔離が実施され、陽性患者への接触者は2+12日間隔離される。（中略）今夜12時に封鎖される団地重点区域リストを添付した』」というものがあった。

第3期と第1期では、「疫苗（ワクチン）」「感染」「能（できる）」という語が共通していた。第1期はワクチンがまだ開発されていなかった。事例として、「通常の肺炎予防のためのワクチンで新型コロナウイルスの感染を防ぐことができる」というフェイクニュースがあった。そして、第3期には、既にワクチンの普及に伴い、人々はワクチンの副作用や効果にも関心を持っていた。「最近、ネット上で広範囲に流れているチャットの記録によれば、多くの人々が健康診断で肺の結節（しこり）が確認されたが、ワクチンを接種する前は結節がなかった。肺の結節はワクチンの接種によって引き起こされる」というものがあった。

4.5 「期間」を外部変数とした対応分析結果と考察

このKH Coderを開発した樋口（2023）によれば、対応分析は、外部変数を原点（0,0）に近い位置にある語は外部変数に関係なく満遍なく出現しているため、特徴がない語である。逆に、原点から離れている語ほど、特徴的な語であるとしている。

「期間」を外部変数として対応分析を行った結果が、図-6である。3期間の布置を概観すると、第1期と第2期は近い位置にあり、第3期は離れている。この点から、第1期と第2期の内容に含まれる語は、共通しているものが多かったことがわかる。

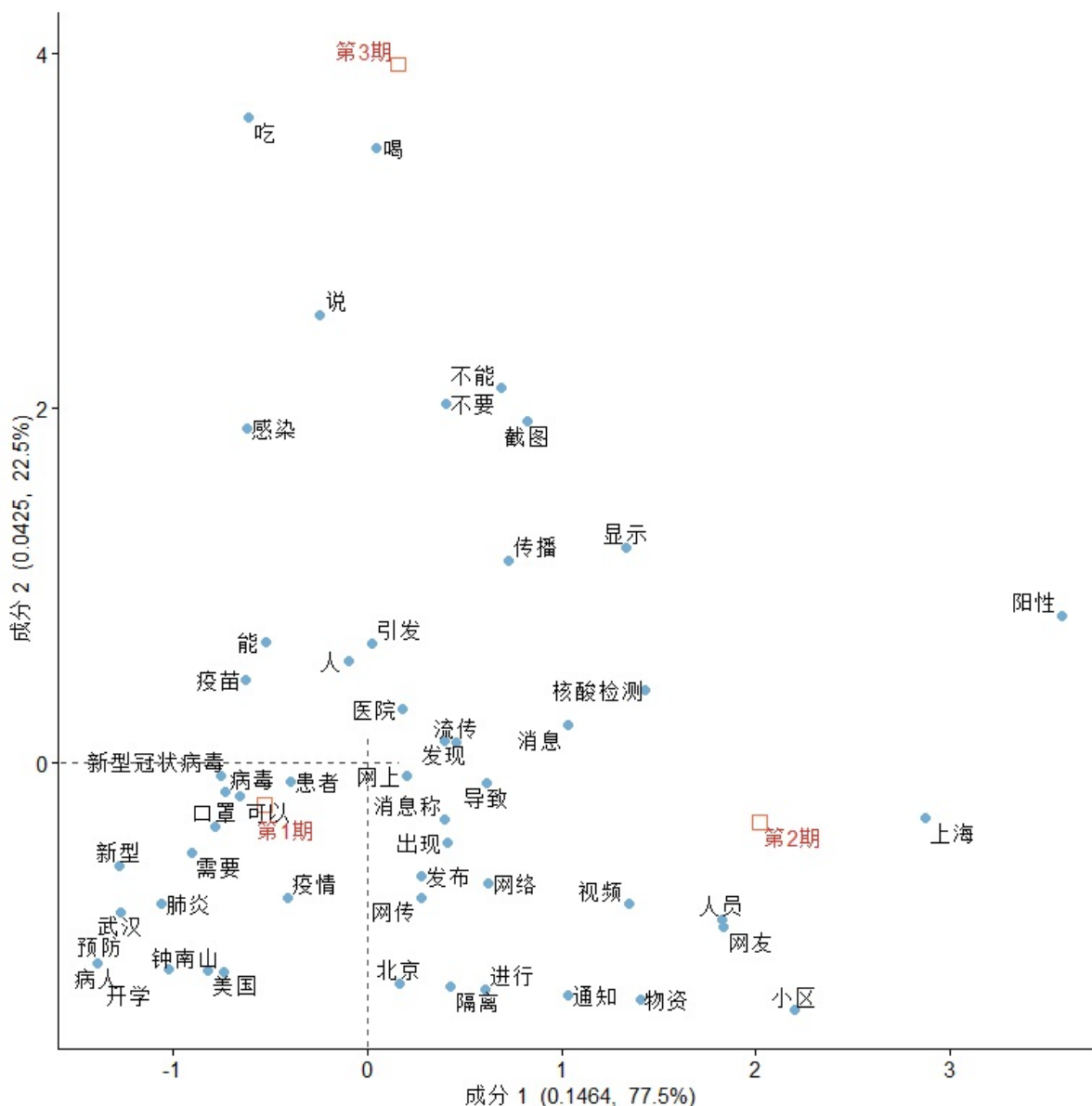


図-6 「期間」を外部変数とした対応分析の結果

第1期に特徴的な語は、原点より離れた左下に固まっている。「武汉（武漢）」「新型」「病人」「预防（予防）」「开学（学校が始まる）」「钟南山（鍾南山）」「美国（米国）」などである。武漢市で発生した新型コロナウイルス感染症は、当初、予防やワクチンなどの医療公衆衛生情報が不明確だったため、大量のフェイクニュースが発生した。「鍾南山」とは、新型コロナウイルス感染症に関する中国政府の責任者（医師）であり、彼の名前は情報の権威付けとして使われた。また、「学校が始まる」は、学校が閉鎖されるのか始まるのかどうか注目が集まっており、政府や自治体政府による公式発表「告知」として、フェイクニュースが流れたことが特徴であった。

この第1期におけるフェイクニュースの事例として、「長沙にある中南大学湘雅病院はテントを設置して、武漢から来る肺炎患者を受け入れる」「イチゴを多く摂取することは、肺炎の予防に大きな効果がある」などというものがあった。

次に、第2期に特徴的な語は、右下に現れている。「上海」「小区（団地）」などである。この時期は、上海の団地でロックダウンが行われた時期であり、この2つの語が特徴的な語になったのは必然的であろう。この時期の事例として、「今日（25日）の朝から、ネット上で『現在、上海には上海ナンバー以外の車は入れない』という情

報が流れている」「最近、上海・嘉定のある団地住民が、物資が供給されないために飛び降りた、という動画がネット上で流れている」などというものがあつた。

最後に、第3期に特徴的な語は、真上に現れている。「喝（飲む）」「吃（食べる）」である。第3期は、一旦落ち着いた新型コロナウイルス感染症の新規患者数が、再度、複数の都市で増加した時期である。このため、第1期と同じように、「飲む」「食べる」ことによって予防できるというフェイクニュースが発生したと考えられる。この時期における事例として、「電解質水は新型コロナの症状を緩和させることができる。新型コロナに感染した場合、白湯を飲むのは効果がなく、電解質水を飲むべきだ」「38.5°C以下の発熱の場合は、解熱剤を摂取してはいけない。38.5°C以上の場合のみ解熱剤を摂取して良い」などがあつた。

5. 結論

5.1 フェイクニュースの全体的特徴

本研究の新規性・独自性は、新型コロナウイルス感染症の発生から終息までの3年間において発生したフェイクニュースを研究対象にしたこと、テキストマイニング分析によって全体構造、3期間ごとの特徴、時系列的変化を明らかにしたことである。最後に、本研究の成果について簡潔にまとめる。

本論文は最初に、3つの目的を設定した。

1つ目の目的は、微信「較真」によってフェイクニュースと判定された3年間の情報を、テキストマイニングツールKH Coderを使って、全体構造を明らかにすること。

これについては、まず頻出語リストの分析によって、以下のことが明らかになった。

まずフェイクニュースに含まれる頻出語のリストを作成したところ、5つの大きな語群（グループ）があることがわかつた。①「新型コロナウイルス」そのものに関連するグループ、②情報源・伝達経路に関連するグループ、③拡散メディアに関連するグループ、④場所や地域に関連するグループ、⑤予防法に関連するグループ、である。

武漢で発生した新型コロナウイルスは、未知のものであつたために、当初は、ウイルスそのものに関するフェイクニュースが多く発生した。それらは、ネット上で拡散し、主に「動画」を視聴することで、流布されていった。また、武漢市に続いて、上海市でも感染が爆発し、ロックダウンによって団地が封鎖された。このため、閉じ込められた人たちが自殺したり飛び降りたりしようとした、などのフェイクニュースが発生した。全体的に、予防法に関するフェイクニュースが大量に発生し拡散した。

共起ネットワーク分析によって、フェイクニュースには3つの大きなグループが存在し、関連しあつていることがわかつた。1つ目は「新型コロナウイルス・予防法」に関するもの、2つ目は「食べる・陽性・できない」という日常生活上の禁止・許可事項、3つ目は「流布・ネット・動画」など拡散経路に関するものであつた。つまり、「新型コロナウイルス」の実体や「予防法」に関連するもの、日常生活上の禁止・許可事項に関するフェイクニュースが多数発生し、それらはネット上で「動画」を視聴することで拡散していったことがわかつた。

さらに、独立したグループには、「ワクチン」の副作用に関するもの、「米国—中国」の関係に関するもの、「モーメンツ（友達グループ）」に関するものがあつた。特に、「モーメンツ」は、微信のグループチャットであり密接なコミュニケーションが行われるため、フェイクニュースの拡散速度が増したと考えられる。

5.2 3期間ごとのフェイクニュースの特徴と相互関係

2つ目の目的は、KH Coderによる外部変数を使った分析で、3期間のフェイクニュースの時系列的変化を明らかにすること。「期間」を外部変数として共起ネットワーク図の作成を試みた。

第1期では「新型」「武漢」「ウイルス」「肺炎」の単語が特徴的であつた。第1期は発生当初のことであり、ロックダウンされた武漢市、そして未知のウイルスそのものに関するフェイクニュースが多かつた。

第2期は、上海市での感染の爆発により、都市がロックダウンされた。頻出語は「上海」「団地」「PCR検査」「告知」「物資」などである。特に「ロックダウンされた団地から飛び降りて脱出しようとした」などのフェイクニュースが発生した。

第3期では、複数の都市で無症状の感染者や再感染者が増加したため、第1期と同じように「予防法」に関するフェイクニュースが再び増加した。

以上のことから、第1期は「武漢」「新型コロナウイルス」に関連するフェイクニュース、第2期は「上海」「団地」の語を含むフェイクニュースが多数発生したことがわかつた。第3期は、ワクチン接種が進んできたものの、各都市で新規患者が増加し、第1期と同じ「予防法」に関するフェイクニュースが発生したことが明らかになつた。

た。

5.3 3 期間ごとの特徴的な語群

3 つ目の目的は、KH Coder の対応分析を使用して、3 期間ごとの特徴的な語群を割り出し、その語群から各期間のフェイクニュースの特徴を明らかにすること。

その結果、第 1 期は発生した場所である「武漢」や未知の「新型コロナウイルス」そのものに関するものに加え、「学校が始まる」など政府機関の公式発表「告知」を装ったフェイクニュースが多数発生したことが特徴であった。第 2 期は、ロックダウンされた「上海」「団地」に関するフェイクニュースが発生したことが特徴であった。第 3 期は、第 1 期と同じように、「飲む」「食べる」と関連した「予防法」に関するフェイクニュースが多数発生したことが特徴であった。時間的な経過によって、「新型コロナウイルス」「ワクチン」に関する知識が普及したものの、第 3 期には、再び複数の都市で感染が広がったことから、「予防法」に関するフェイクニュースが発生したと考えられる。

総括すると、パンデミックが発生した場合は、最初に、ウイルスそのもの、予防法に関するフェイクニュースが多数発生すること。特に、予防法に関するものは、その後も繰り返し発生すること。フェイクニュースは主に、ネット上の「動画」で拡散していくこと。また、ロックダウンのような厳しい政策をとった場合は、公的機関になりすました「告知」に関するフェイクニュースが発生すること。このため、政府や自治体は、フェイクニュースについては打ち消し報道（広報）をすると同時に、積極的に正確な情報を公開する必要があること。以上のことを、本研究の結果から考察した。

[注]

- 1 「騰訊公布二零二二年第四季及全年業績（テンセント社が発表した 2022 年第 4 四半期および年間業績）」によれば、2022 年の月間アクティブユーザー数は 13.13 億人である。
- 2 微博（ウェイボー）とは中国における代表的な SNS の 1 つで、ミニブログサイトのような形式である。微博データセンターの統計データによれば、2023 年 6 月までの微博の月間利用者数は 5.99 億人である。
- 3 知乎（ジーフー）とは中国における代表的な SNS の 1 つで、Q&A サイトのような形式である。知乎 2023 年 6 月までの決算によれば、2023 年第 2 四半期の平均月間利用者数は 1.06 億人である。

[参考文献]

- [1] 坂本 旬「偽情報時代のメディア情報リテラシーと学校図書館」『メディア情報リテラシー研究 = The Journal of Media and Information Literacy』, 第 1 巻, 第 2 号, pp.83-100, (2020).
- [2] 笹原 和俊『フェイクニュースを科学する：拡散するデマ、陰謀論、プロパガンダのしくみ』化学同人, (2018).
- [3] 笹原 和俊「ウェブの功罪」『情報の科学と技術』, 第 70 巻, 第 6 号, pp.309-314, (2020).
- [4] 白崎 護「フェイクニュースとメディア環境」『研究論集』, 第 112 巻, pp.331-349, (2020).
- [5] 末吉 美喜『テキストマイニング入門：Excel と KH Coder でわかるデータ分析』オーム社, (2019).
- [6] 立岩 陽一郎, 楊井 人文『ファクトチェックとは何か』岩波書店, (2018).
- [7] 中森 広道「災害流言の展開とその特性」『消防防災の科学』, 第 139 号, pp.34-39, (2020).
- [8] 樋口 耕一「計量テキスト分析における対応分析の活用」『コンピュータ&エデュケーション』, 第 47 巻, 第 0 号, pp.18-24, (2019).
- [9] 樋口 耕一『社会調査のための計量テキスト分析：内容分析の継承と発展を目指して』ナカニシヤ出版, (2020).
- [10] 樋口 耕一, 中村 康則, 周 景龍『動かして学ぶ! はじめてのテキストマイニング：フリー・ソフトウェアを用いた自由記述の計量テキスト分析』ナカニシヤ出版, (2022).
- [11] 福長 秀彦「調査研究ノート SNS 時代の誤情報・虚偽情報とマスメディアの打ち消し報道; 留意すべき事柄を考える」『放送研究と調査』, 第 69 巻, 第 8 号, pp.100-110, (2019).
- [12] 福長 秀彦「新型コロナウイルス感染拡大と流言・トイレトペーパー買いだめ」『放送研究と調査』, 第 70 巻, 第 7 号, pp.2-24, (2020).
- [13] 李 瑞鵬, 松野 良一「中国における新型コロナウイルス感染症に関するフェイクニュースの分類と分析：微信「較真」がファクトチェックした結果の分析を中心に」『情報文化学会誌』, 第 28 巻, 第 2 号, pp.19-26, (2022).

- [14] 李 瑞鵬, 松野 良一「中国における新型コロナウイルス感染症に関するフェイクニュース—時系列的変容を中心に—」『情報社会学会誌』, 第 18 卷, 第 1 号, pp.49-61, (2023).
- [15] 程 曾芬「騰訊較真平台事事实查証新聞生産机制研究」, 中央民族大学ニュースと伝播研究科修士論文, p. 27, (2018).
- [16] 国家信息中心・南京大学網絡伝播研究院『2020“新型冠状病毒肺炎” 公众認知与信息伝播調研』, pp. 07-20, (2020).
- [17] 古麗拌合熱姆 巴图爾・田瑞・郭蕊「新冠病毒疫情相關謠言特点分析—基于較真查証平台的研究」『中国社会医学雜誌』, (02), pp. 141-145, (2023).
- [18] 黄 佩英「从謠言心理学角度分析“塩慌”産生的根源」『語言文学』, 03, p. 220, (2011).
- [19] 刘 鹏程・孫 梅・李 程躍・張 建華・宁宁・応 曉華・黄 葭燕・陳 文・郝 模, 「H7N9 事件網絡輿情分析及其对突發公共衛生事件应对的启示」『中国衛生事業管理』, 10, p. 785, (2014).
- [20] 錢 迎飛「新冠肺炎疫情影响期間的謠言研究—基于“較真”辟謠平台 437 条樣本的分析」『視聽』08, pp. 132-133, (2022).
- [21] 騰訊公布二零二二年第四季及全年業績,
<https://static.www.tencent.com/uploads/2023/03/22/8d304603bd688ad947575f738ce62fca.pdf> (2023 年 10 月 12 日アクセス).
- [22] 王 丹「騰訊新聞“較真”欄目研究」『西部学刊』, 05, pp. 56-57, (2016).
- [23] 吴 世文「重大突發公共衛生事件中的偽信息伝播, 治理困境及其突破路径——以新冠肺炎疫情為例」『電子政務 E-GOVERNMENT』, 09, p. 48, (2020).
- [24] 張 皓月「突發公共衛生事件網絡謠言治理研究—以新型冠状病毒肺炎疫情影响為例」, 長春工業大学ニュースと伝播研究科修士論文, pp. 23-24, (2020).
- [25] 中国互聯網情報中心 (CNNIC) 『第 45 回中国互聯網發展狀況統計報告』CNNIC 出版, pp. 34-35, (2020).
- [26] Allport Gordon W. (Gordon Willard), Postman Leo Joseph, 南 博『デマの心理学』岩波書店, (2008).
- [27] Soroush Vosoughi, Deb Roy, Sinan Aral, The spread of true and false news online, Science, Vol 359, p. 1150, (2018).
- [28] 中国衛生健康委員会「説明」,
<http://www.nhc.gov.cn/xcs/yqtb/202212/7272431ee60c4f4d953b0c16257c230e.shtml> (2023 年 8 月 12 日アクセス).
- [29] 中国疾病予防管理センター「全国新型コロナウイルス感染症感染状況」,
https://www.chinacdc.cn/jkzt/crb/zl/szkb_11803/jszl_13141/202304/t20230429_265709.html (2023 年 8 月 12 日アクセス).

iqq690176825@yahoo.co.jp

(2024 年 3 月 30 日受理)