

# Contagious words and epidemic behaviors: A usage-based exploration of internet neologisms related to COVID-19

Ms. Siyu Lei

PolyU/XJTU

Prof. Ruiying Yang

XJTU

Prof. Chu-Ren Huang

PolyU/PolyU-PKU

17/10/2020

# Outline

- Overview of the research
- Research Design
- Results
- Discussion



# Overview of the research

**Background:** Internet neologisms are contagious and reflect collective human behavior. Can this observation be leveraged to reflect an epidemic situation by tracking the development of internet neologisms over time?

**Significance:** This paper correlates the competition of neologisms with the development of an epidemic as collective human behavior changes, enhancing our understanding of these two phenomena.

**This study:** Specifically, this study tracks the use of **COVID-19 neologisms from late December 2019 to the end of April, 2020** based on **Baidu Index**.

# Overview of the research

## Research questions:

- (1) **What different COVID-19 neologisms** are used for search over internet?  
How can these neologisms be **categorized**?
- (2) Do usage variations of COVID-19 neologisms over time reflect the **impact of language-internal factors** as well as **social psychological situation** as the **changes of collective human behaviors**?
- (3) Is there any **correlation** between neologism uses and pandemic development (i.e., new confirmed cases, new suspected cases, new deaths, and current suspects)?

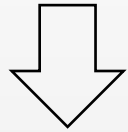
# Overview of the research

## Results:

- (1) The neologisms are designated by **five categories**: under-specified references, pre-official names, pejorative names, official names, and English abbreviations.
- (2) Language-internal factors (i.e., **frequency**) and the changes of social psychological situation (i.e., **policy and emotion**) have impact on lexical competition and evolution.
- (3) There is a **strong correlation** between neologism uses and pandemic development.

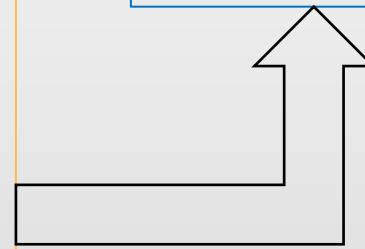
# Research Design

- Step 1. --- RQ 1.
  - Baidu: COVID-19 variants
  - Categorization



- Step 2. --- RQ 2.
  - Baidu Index: Frequency
  - Time range: 5 days per search
  - Influential factors: language-internal factors, i.e., frequency; social psychological factors, i.e., policies and emotions

- Step 3. --- RQ 3.
  - COVID-19 pandemic website: the number of new confirmed cases, new suspected cases, new deaths, and existing suspected cases at different times



# Results

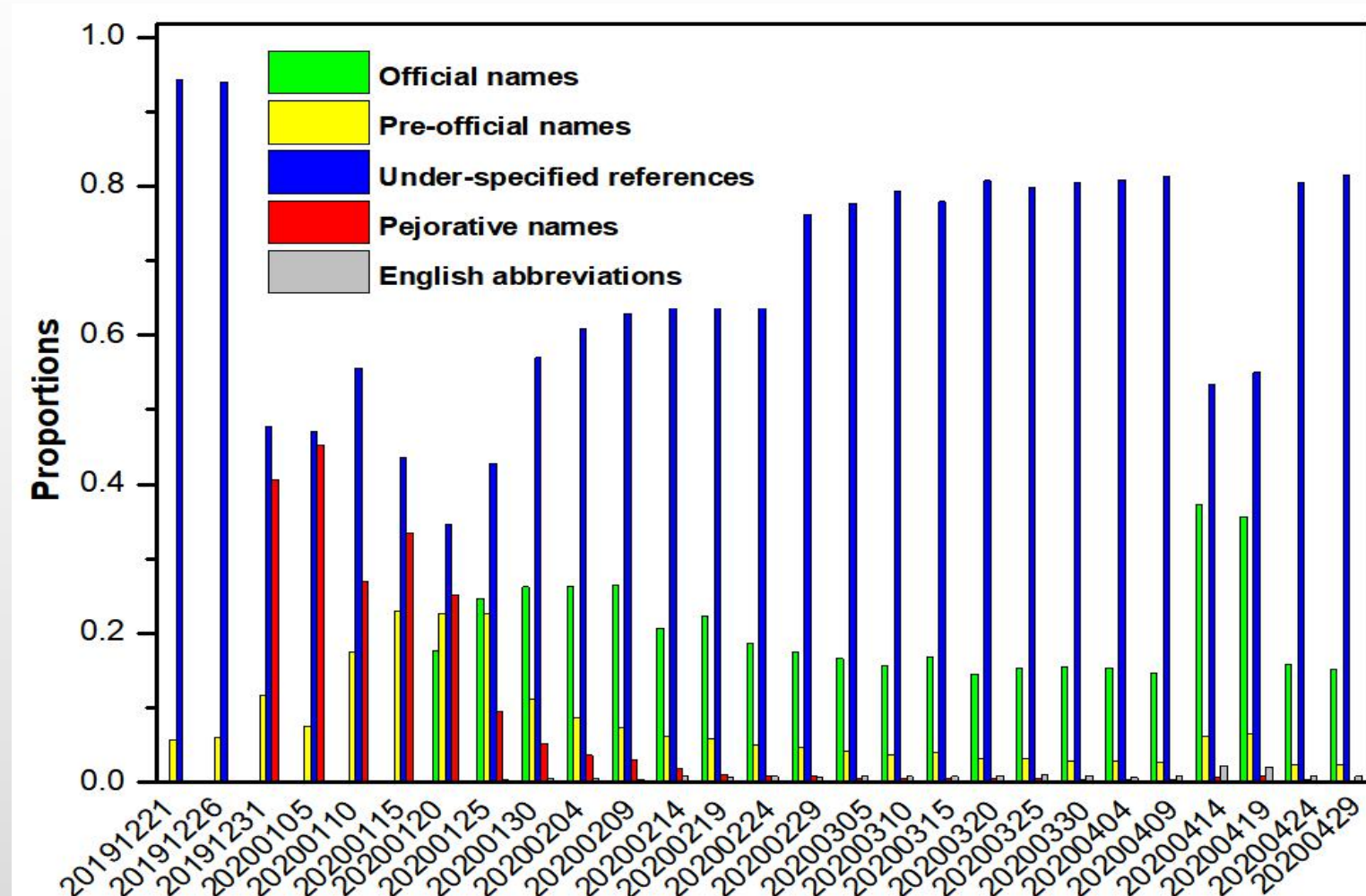
## RQ 1. COVID-19 variants and categories

Table 1 Different categories and their variants of COVID-19 neologisms

Categories	Variants
Official names	新型冠状病毒肺炎 ‘novel type corona shape virus pneumonia’
	新冠肺炎 ‘novel corona pneumonia’
	新冠疫情 ‘novel corona epidemic’
	新冠病毒 ‘novel corona virus’
	新冠 ‘novel corona’
	新型冠状病毒 ‘novel type corona virus’
	2019 新型冠状病毒 ‘2019 novel type corona shape virus’
Pre-official names	不明原因肺炎 ‘pneumonia of unknown reasons’
	病毒性肺炎 ‘viral pneumonia’
	新型病毒 ‘novel type virus’
	新型肺炎 ‘novel type pneumonia’
	冠状病毒 ‘corona shape virus’
Under-specified references	疫情 ‘epidemic’
	肺炎 ‘pneumonia’
Pejorative names	武汉肺炎 ‘Wuhan pneumonia’
	武汉病毒性肺炎 ‘Wuhan viral pneumonia’
	中国病毒 ‘China virus’
	武汉新型肺炎 ‘Wuhan novel type pneumonia’
	武汉病毒 ‘Wuhan virus’
English abbreviations	COVID-19
	2019-nCov
	<u>Coronavirus</u>
	SARS-CoV-2
	<u>nCov</u>

# Results

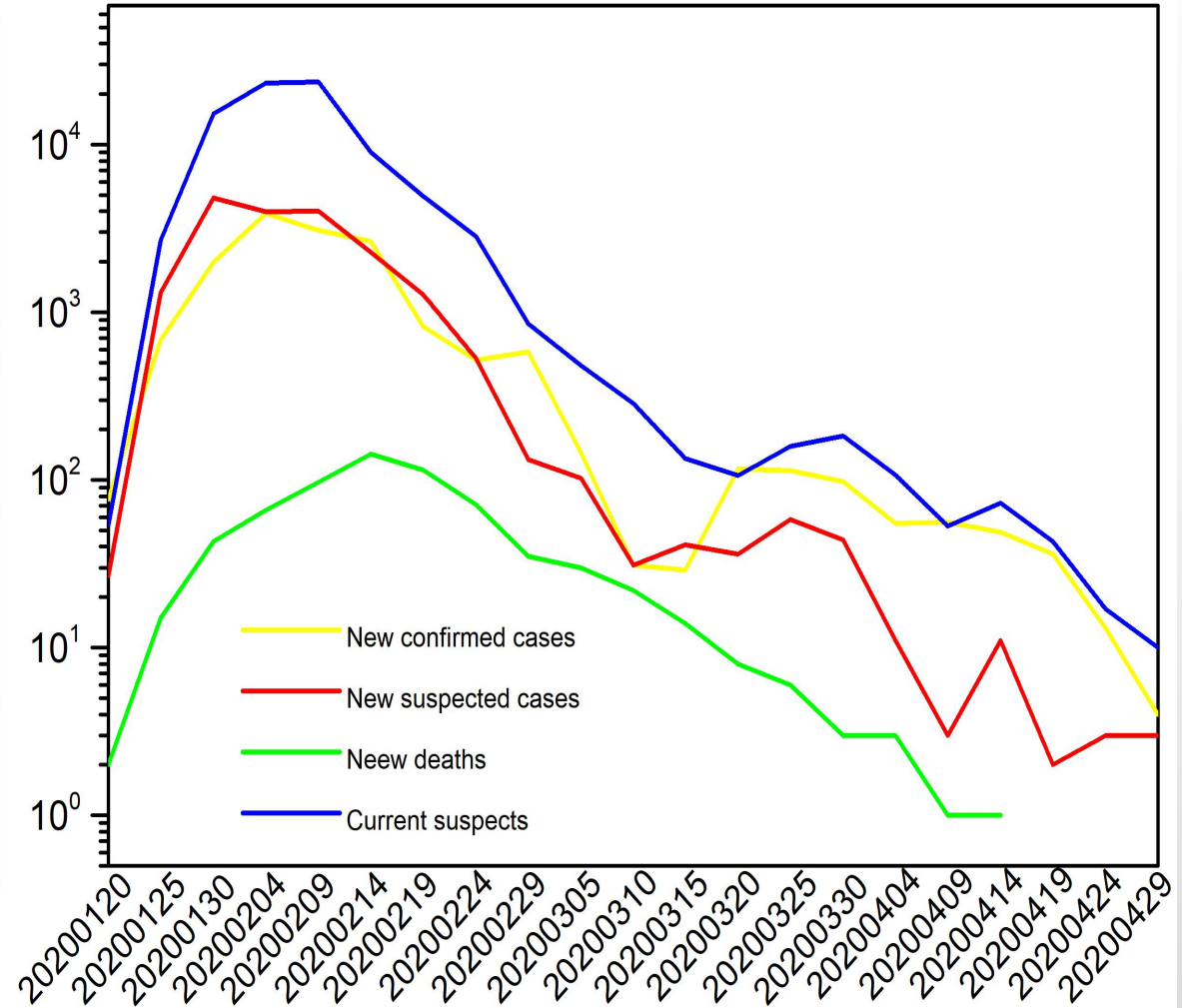
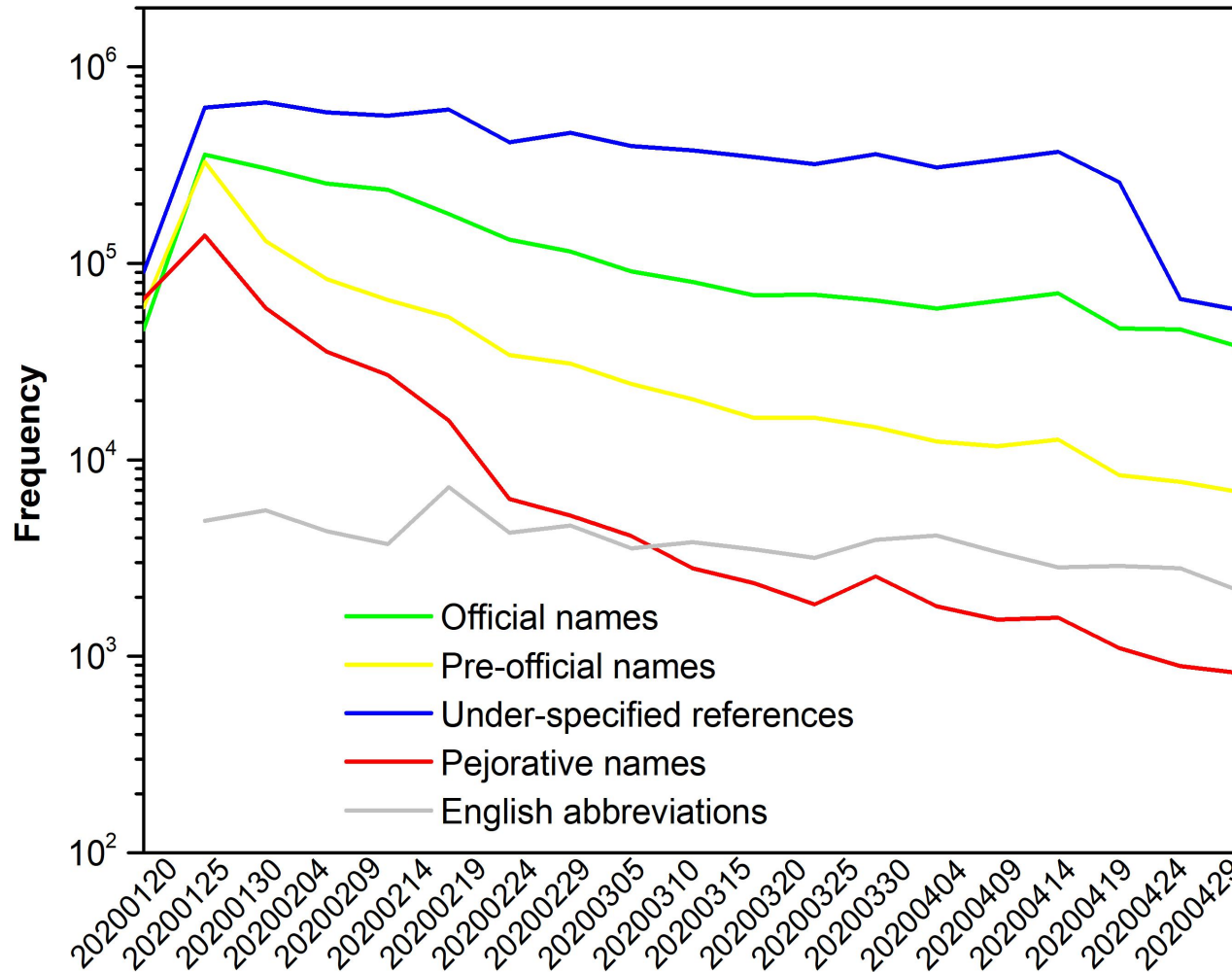
RQ 2. Impact of frequency, policies, and emotion





# Results

RQ 3. Correlation of internet neologisms and pandemic



# Results

RQ 3. Correlation of internet neologisms and pandemic

Table 2 Correlation analysis between COVID-19 neologisms and pandemic data

Word categories	New confirmed cases	New suspected cases	New deaths	Current suspects
Official names	.725**	.828**	.492*	.715**
<u>Pre-official names</u>	.336	.448	.132	.297
Under-specific references	.715**	.749**	.651**	.662**
Pejorative names	.288	.416	.047	.274
English abbreviations	.555*	.545*	.683**	.416

\* $p < .05$ ; \*\* $p < .01$

# Discussion

- According to the principle of disease nomenclatures in 17th century of Western Europe, there are 5 categories to summarize different COVID-19 variants in our research.
- In addition to the most typical factor, i.e., frequency, that can influence the competition of different categories over time, policies designed by WHO and National Health Center of China are also an indispensable factor. Meanwhile, emotional changes also explain such competition.
- Quantitative results showed there is a correlation between COVID-19 neologisms and pandemic data.
- We thus summarized at what stage of COVID-19 would cause what emotional situation and then which word categories would be more likely to be used at that time in a flow chart.
- We also described the development of COVID-19 neologisms in an inverse N-curve.

Thank you so much!

Any comments and suggestions?