# 정치적 억압으로 훼손된 소수집단의 명성 회복 조건 연구

20171218 김다빈

## 1. 정치적 억압으로 훼손된 소수집단의 명성 회복 조건

정치적 억압으로 훼손된 소수집단의 명성은 어떤 조건에서 회복되는가. 알려진 사실에 따르면 비미활동위원회가 블랙리스트로 지정한 할리우드 종사자 10인은 1940년대 말에 영화산업에서 배제되었으며, 1950년대 조지프 매카시 상원의원의 몰락을 계기로 명예를 회복하기 시작했다. 나아가 1959년 트루먼 전 대통령의 위원회 비판을 계기로 블랙리스트의 의미가 약화되면서 영화계에 본격적으로 복귀하였다. 하지만, 이 사실만으로 꾸준히 지지를 얻었던 조지프 메카시의 공산주의자 색출작업에 대한 평가가 반전된 계기는 명확히 확인하기 어렵다. 당시 미국사회는 공화당과 민주당은 모두 공산주의자 낙인을 피하기 위해 조지프 메카시의 행보에 반론을 제기하지 않았으며, 할리우드 영화계 또한 그의 발언에 동조하고 힘을 실었다. 이러한 배경에서 조지프 매카시에 대한 평가가 하락하고, 나아가 트루먼 전대통령이 공개적으로 위원회를 비판한 사실은 당시 미국사회의 사회적 정치적 분위기에 근본적인 변화가 나타났을 가능성을 암시한다. 필자는 할리우드 10인과 조지프 매카시의 명성에 변화를 불러온 60년대 전후 미국사회의 변화에서 '정치적 억압의 피해자가 명성을 회복하기 위한 조건'을 발견할 수 있을 것이라고 판단했으며, 'US\_suppression.txt' 텍스트 데이터와 Google Ngrams 함수를 활용해 조사를 진행하였다.

#### \* 들어가기 전 : 일반적으로 알려진 조지프 매카시 명성 하락의 계기

2차 세계대전 종료 후 본격적인 미소 체제 대립이 시작된 1940년대 말, 미국사회에는 소련 간첩과 공산주의에 대한 두려움이 팽배하였다. 이러한 배경에서 공화당 상원의원 조지프 매카시는 미국 내에 존재를 감춘 공산주의자가 활동하고 있으며, 그들의 명단을 본인이 가지고 있다고 주장하였다. 그의 고발은 충분한 사실관계 파악 없이 무분별하게 진행되었음에도 불구하고 많은 대중의 관심을 받았고, 언론과 정치계 또한 그의 발언에 동조했다.

그러나 1953년 민주당을 이기고 다수당으로 집권한 공화당은 더이상 매카시의 사상적 공세가 불필요하다고 판단했다. 결국 공화당 의원의 양심선언을 계기로 매카시즘에 대한 비판이 시작되었으며, 1954년 랠프 플랜더스 상원의원의 비난 결의안이 채택되면서 매카시의 영향력이 막을 내렸다.

해당 정보에 따르면 정치적 억압의 피해자들은 '집권당 변화' 혹은 '당원의 양심선언' 등 단편적인 사건을 계기로 회복되었다. 하지만 필자의 2019년 미국영어 코퍼스 데이터 조사결과에 따르면 억압집단의 명예 회복은 위 사건에 따라 즉각적으로 나타나지 않았으며, 보다 장기적인 사회적 변화 기류에 따라진행되었다.

# 2. Hollywood Blacklist 데이터 특성 확인

조사를 위해 빅데이터학 3차과제 첨부파일 'US\_suppression.txt'를 ndarray 데이터로 불러와 list로 변환하였다. 이후 Google Ngrams 함수를 활용해 1900년부터 2000년까지 'eng\_us\_2019' 코퍼스에 나타난 할리우드 블랙리스트 10인의 빈도 추세 정보를 확보하였다. 해당 데이터는 영화인의 이름 Columns와 연도 Index로 이루어진 DataFrame으로, 이에 대한 라인차트 시각자료는 다음과 같다.

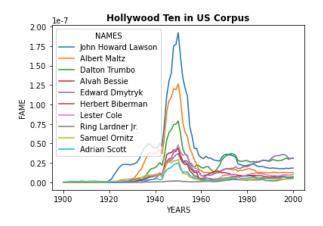
### # US\_suppression.txt 데이터 불러오기

```
*matplotlib inline
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np
from GoogleNgrams import GoogleNgrams
US_suppression = list(np.loadtxt('US_suppression.txt', dtype='str', delimiter='\n'))
US_suppression
['John Howard Lawson',
 'Albert Maltz',
 'Dalton Trumbo',
 'Alvah Bessie',
 'Edward Dmytryk'
 'Herbert Biberman',
 'Lester Cole',
 'Ring Lardner Jr.',
 'Samuel Ornitz',
 'Adrian Scott']
```

#### # 할리우드 블랙리스트 10인의 빈도 변화

```
US_sup = GoogleNgrams(', '.join(US_suppression), B = 1900, C=2000, D='eng_us_2019')
US_sup.columns.name = 'NAMES'
US_sup.index.name = 'YEARS'
US_sup.plot()
plt.ylabel('FAME')
plt.title('Hollywood Ten in US Corpus', weight='bold')
```

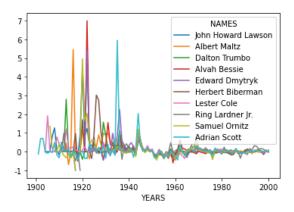
Text(0.5, 1.0, 'Hollywood Ten in US Corpus')



### # 할리우드 블랙리스트 10인의 변동률 (비정규분포)

```
US_sup.pct_change().plot()
```

#### <AxesSubplot:xlabel='YEARS'>



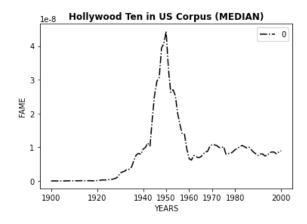
위 자료에 따르면 할리우드 10인의 이름 빈도는 1940년대에 급격한 상승세를 보이며 최고점에 도달하나, 약 1950년을 기점으로 급격히 하락한다. 1960년부터는 다시 빈도가 서서히 증가하며, 1970년대 들어서 일정 수준 이상의 값을 회복한다.

필자는 해당 데이터의 표본 크기가 작고 정규분포와 거리가 멀어 산술평균을 적용하면 왜곡된 결과를 얻을 것이라고 판단했다. 따라서 할리우드 블랙리스트 인물의 명성 추세를 중앙값으로 단순화했으며, 추세의 전환지점을 포착해 '억압구간'과 '회복구간'을 정의하였다. 다음은 line 차트와 cumsum line 차트로 살펴본 중앙값 추세 시각자료다.

# median 메서드를 적용한 중앙값 라인차트

```
US_sup_MED = pd.DataFrame(US_sup.median(1))
US_sup_MED.plot(linestyle='-.', color='k')
plt.ylabel('FAME')
plt.xticks([1900,1920,1940,1950,1960,1970,1980,2000])
plt.title('Hollywood Ten in US Corpus (MEDIAN)', weight='bold')
```

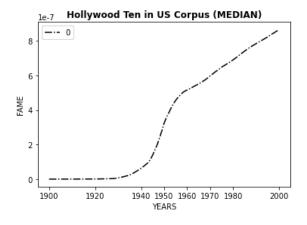
Text(0.5, 1.0, 'Hollywood Ten in US Corpus (MEDIAN)')



# cumsum() 메서드를 적용한 누적합계 중앙값 라인차트

```
US_sup_MED.cumsum().plot(linestyle='-.', color='k')
plt.ylabel('FAME')
plt.xticks([1900,1920,1940,1950,1960,1970,1980,2000])
plt.title('Hollywood Ten in US Corpus (MEDIAN)', weight='bold')
```

Text(0.5, 1.0, 'Hollywood Ten in US Corpus (MEDIAN)')



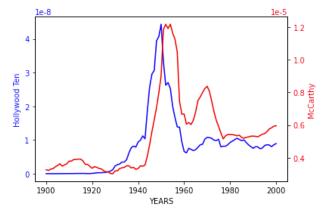
위 두가지 라인차트에 따르면 할리우드 블랙리스트 10인의 명성은 1950년부터 1960년 사이에 극적으로 하락한다. 1960년부터 1970년 사이에는 점전직으로 명성을 회복하며, 이후로는 안정적으로 영향력을 유지하는 모습을 보인다. 이에 따라 필자는 1950년부터 1960년까지를 '억압 기간', 1960년부터 1970년까지를 '회복 기간'으로 정의하고 명성회복의 조건을 탐색하였다.

## 2. 조건 1 : 조지프 매카시(Joseph McCarthy)의 명성변화

```
US_sup_MED = pd.DataFrame(US_sup.median(1))
US_McCarthy = GoogleNgrams('McCarthy', B = 1900, C=2000, D='eng_us_2019')

fig, ax1 = plt.subplots()
ax1.set_xlabel('YEARS')
ax1.set_ylabel('Hollywood Ten', color='blue')
ax1.plot(US_sup_MED, color='blue')
ax1.tick_params(axis='y', labelcolor='blue')

ax2 = ax1.twinx()
ax2.set_ylabel('McCarthy', color='red')
ax2.plot(US_McCarthy, color='red')
ax2.tick_params(axis='y', labelcolor='red')
```



첫째로 할리우드 10인에 대한 정치적 억압을 주도한 조지프 매카시 상원의원의 명성 추세를 살펴봤다. 위 라인차트는 Google Ngrams에서 나타난 블랙리스트 인물의 빈도 중앙값과 조지프 매카시의 빈도 추세를 비교한 자료이다. 이에 따르면 할리우드 10인의 명성은 조지프 매카시의 명성이 최고점에 이르기 직전에 급격히 하락했으며, 그가 1950년대 초반까지 굳건히 명성을 유지하는 동안 지속적으로 영향력을 잃었다. 이 자료에서 주목할만한 점은 1950년대에 매카시가 영향력을 잃고 1957년에 사망한 후로도 블랙리스트 인물의 명성 하락이 지속되었다는 점이다. 이들의 빈도가 본격적으로 상승추세에 접어든 시점은 메카시의 죽음 후 3년이 지난 1960년대로 나타났다. 이러한 결과는 한명의 정치적 인물의 명성하락이 정치적 억압의 피해자들에 대한 명성 회복으로 직결되지 않는다는 사실을 보여준다. 나아가 정치적 억압을 주도한 개인이 사망하더라도, 그 잔재가 지속되어 기존의 정치적 억압이 지속될 수 있음을 암시한다.

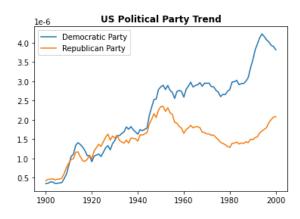
## 3. 조건2: 정당의 명성 변화

앞서 살펴본 자료를 통해 '개인'의 명성 변화보다 '정치집단'의 명성 변화를 확인하는 것이 더 바람직하다고 판단했다. 미국은 민주당과 공화당이 높은 정치적 영향력을 행사하는 대표적인 양당제 국가로서, 두 정당의 명성 변화를 통해 정치적 억압집단의 명성 회복 조건을 발견할 수 있을 것이라고 추측했다.

역사적 사실로는 1933년부터 1953년까지 민주당이 집권당을 유지하고 1953년부터 1961년까지 공화당이 집권당을 차지했지만, 공화당 상원의원인 조지프 매카시의 명성이 1953년 이전에 최고점을 유지하고 1953년 이후 하락한 점을 고려했을 때, '집권당' 여부가 당의 높은 영향력을 대변하지는 않는다고 판단했다. 필자는 두 당의 실질적인 명성을 파악하기 위해 '민주당'과 '공화당'을 가리키는 공식명칭인 'Democratic Party'와 'Republican Party'를 Google Ngrams 에 입력했다.

### # 비교 1: Democratic Party, Republican Party

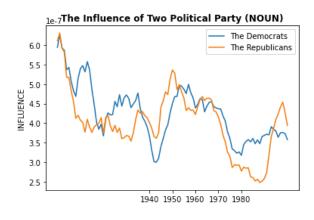
Text(0.5, 1.0, 'US Political Party Trend')

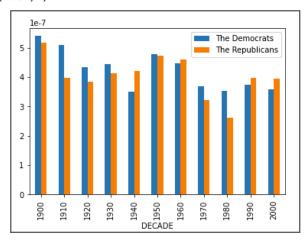


할리우드 10인의 억압기간과 회복기간에 해당하는 1950년부터 1970년대까지의 양당의 빈도추세 변화를 확인했을 때, 유의미한 결과를 확인하기가 어려웠다. GoogleNgrams에서 제공하는 어휘사용 사례를 확인하니, 대체로 Democratic Party, Republican Party는 함께 사용되는 경우가 많았다.

#비교 2: The Democrats, The Republicans (line chart, bar chart)

Text(0.5, 1.0, 'The Influence of Two Political Party (NOUN)')



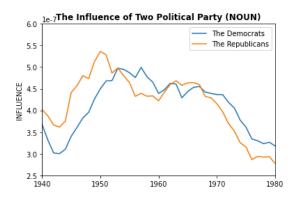


```
US_dem_rep_PARTY = GoogleNgrams('The Democrats, The Republicans',

B = 1900, C=2000, D='eng_us_2019')

US_dem_rep_PARTY.plot()
plt.axis([1940,1980,2.5e-7,6e-7])
plt.xticks([1940,1950,1960,1970,1980])
plt.ylabel('INFLUENCE')
plt.title('The Influence of Two Political Party (NOUN)', weight='bold')
```

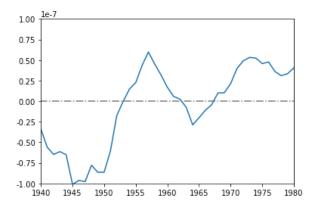
Text(0.5, 1.0, 'The Influence of Two Political Party (NOUN)')



보다 일반적인 서술체에서 등장하는 The Democrats, The Republicans에 대한 GoogleNgrams 데이터를 추출했다. 예상했던 것과 달리 민주당이 집권한 1940~1953년에는 공화당이 민주당보다 높은 빈도값을 보였으며, 반대로 공화당이 집권한 1953~1961년에는 민주당이 공화당보다 높은 빈도값을 나타냈다. 다음은 민주당의 빈도수에서 공화당의 빈도수를 뺀 값의 추세 그래프다.

```
a = US_dem_rep_PARTY['The Democrats'] - US_dem_rep_PARTY['The Republicans']
plt.hlines(0, 1900, 2000, linestyle='-.', color='grey')
plt.axis([1940,1980,-1e-7,1e-7])
a.plot()
```

#### <AxesSubplot:>



해당 차트의 y축을 기준으로 0 이상에 분포한 값은 민주당이 공화당보다 높은 빈도로 나타난 상황을 의미하며, 0 이하에 분포한 값은 공화당이 민주당보다 우세한 빈도로 나타난 상황을 나타낸다. 또한 상승 추세선은 민주당의 빈도 상승, 하락 추세선은 공화당의 빈도 상승을 암시한다. 필자는 GoogleNgrams 데이터가 도서관의 서적 내용에 기반한 자료임을 고려하여 두 당의 빈도 수치를 '정당의 실질적인 영향력'으로 간주하였다.

위 추세선에 따르면 할리우드 10인에 대한 억압이 나타난 1950년에는 공화당의 영향력이 막강하다. 이러한 수치결과는 민주당이 집권한 동안에도 실질적인 정치적 흐름을 주도한 것은 공화당이었음을 보여준다. 그러나 1950년대 초부터 민주당의 영향력이 공화당을 넘어섰으며, 이후 지속적으로 -0.25 이

상의 값이 유지된다. 이러한 추세선 양상은 1953년 이후 민주당이 안정적으로 정치적 영향력을 유지하였음을 나타낸다. 1954년 조지프 메카시에 대한 비판 결의가 채택되고, 1959년 임기를 마친 트루먼 전대통령이 비미활동위원회에 대해 공개적으로 비판한 배경에는 1954년 민주당의 명성 회복이 있었음을 추측할 수 있다.

위 시각자료에 기반했을 때, 특정 정치집단에 의해 훼손된 소수집단의 명성은, 대립하는 정치집단의 영향력이 해당 정치집단의 영향력을 약 5년간 지속적으로 넘어설 경우 회복된다는 가설을 세울 수 있다. 단, 이에 대해서는 다른 국가의 다른 정치적 억압 사례를 확인하여 검증하고 일반화하는 작업이 반드시필요하다.

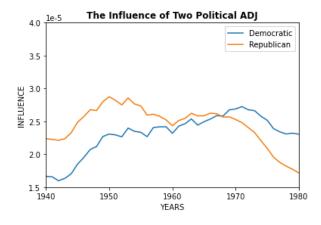
## 4. 조건3: 사회적 관심 (민주당/공화당의 형용사형)

다음으로는 The Democrats와 The Republicans의 형용사형, Democratic과 Republican에 대한 Google Ngrams 데이터를 추출했다. 명사형의 당명 이름이 순수하게 정당 자체를 가리킨 것과 달리, 두 형용사형은 미국사회의 정치적 관심의 흐름을 함께 포착한다. 가령 Democratic 이라는 어휘는 '민주당의'라는 의미로 사용되지만, 동시에 '민주주의의'라는 의미로도 사용된다. Republican 또한 '공화당의'라는 뜻을 지닌 동시에 '공화국의', '공화주의의'라는 의미도 함께 내포한다. 이러한 의미포착의 확대 작업은 '미국 정치집단의 명성 변화'에서 '미국사회의 정치적 의제 변화'로 나아가 보다 거시적인 사회적 변화를 포착하기 위한 시도이다.

#### #비교 3: Democraic, Republican

```
US_dem_rep_ADJ = GoogleNgrams('Democratic, Republican', B = 1900, C=2000, D='eng_us_2019')
US_dem_rep_ADJ.plot()
plt.axis([1940,1980,1.5e-5,4e-5])
plt.xticks([1940,1950,1960,1970,1980])
plt.xlabel('YEARS')
plt.ylabel('INFLUENCE')
plt.title('The Influence of Two Political ADJ', weight='bold')
```

Text(0.5, 1.0, 'The Influence of Two Political ADJ')



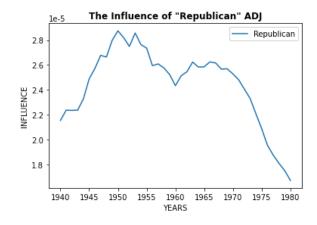
두 형용사의 추세선은 앞서 살펴본 정당명의 결과값보다 안정적이고 일관된 형태를 보인다. 두 형용사는 2차 세계대전이 종료된 이후 1950년까지 함께 급격히 상승하였으며, 1950년부터 1970년까지 'Democatic'은 상승, 'Requblican'은 하강하였다. 그 사이 1960년대 초반에는 두 형용사의 빈도가 비슷하게 나타났으며, 1960년대 후반에 들어서서는 'Democratic'의 빈도가 'Republican'을 월등하게 넘

어섰다. 해당 자료에 따르면 미국사회에 'Requblican'과 연관된 정치적 의제가 주목받은 시기 (1950~1960)에 할리우드 10인에 대한 정치적 억압이 지속되었으며, 반대로 'Democratic'과 연관된 의제가 주목받은 시기(1960~)에 10인에 대한 명성 회복이 이루어졌다.

# 형용사의 미국사회 주요 정치적 의제 반영 (World War II, Communists, National Security, ...)

```
US_rep_ADJ = GoogleNgrams('Republican', B = 1940, C=1980, D='eng_us_2019')
US_rep_ADJ.plot()
plt.xlabel('YEARS')
plt.ylabel('INFLUENCE')
plt.title('The Influence of "Republican" ADJ', weight='bold')
```

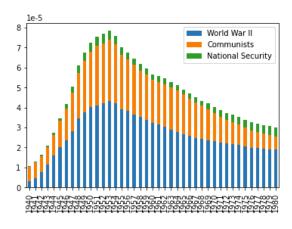
Text(0.5, 1.0, 'The Influence of "Republican" ADJ')



```
SubjectsA = GoogleNgrams('World War II, Communists, National Security',

B = 1900, C=2000, D='eng_us_2019').reindex(range(1940,1981,1))
SubjectsA.plot.bar(stacked=True)
```

#### <AxesSubplot:>



- 미국 정당이름의 형용사형은 해당 집단이 고수하는 정치사상을 드러낸다. 따라서 'Republican'이라는 형용사의 추세도는 개인보다 공동체의 이익을 우선시하는 공화주의적 담론과 밀접한 정치적 주제들의 사회적 관심도를 반영한다. 단, 해당 담론의 세부적인 변화과정을 모두 반영하지는 않으므로, 추후 다양한 형용사 및 의제 빈도조사를 통한 일반화 작업이 필요하다.

## 4. 조건2 & 조건3 비교: 양당 이름의 명사형 - 형용사형

그렇다면 정치적으로 억압된 소수집단의 명성회복을 예측하기 위해서는 The Democratis & The Republicans 명사형과 Democratic & Republican 형용사형 중 어떤 데이터를 확인하는 것이 효과적

일까? 필자는 두가지 추세의 특성을 함께 비교분석하기 위해 두 명사형의 빈도 차(subtract) 'The Democrats - The Republicans'와 두 형용사형의 빈도 차(subtract) 'Democratic - Republican' 그래 프를 하나의 차트에 표시하였다.

# GoogleNgrams 함수로 얻은 두가지 DataFrame (US\_dem\_rep\_PARTY, US\_dem\_rep\_ADJ) 통합

	DEM_NOUN	REP_NOUN	DEM_ADJ	REP_ADJ
1900	5.943880e-07	6.111457e-07	0.000020	0.000039
1901	6.267058e-07	6.308210e-07	0.000021	0.000039
1902	5.926260e-07	5.927591e-07	0.000020	0.000040
1903	5.866815e-07	5.819056e-07	0.000020	0.000039
1904	5.369485e-07	5.179634e-07	0.000019	0.000036

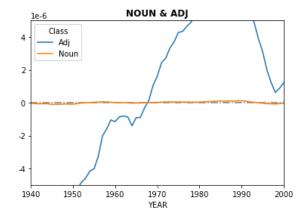
# Hierarchical Indexing 부여를 통한 상위 Column, 하위 Column 설정

Class	Noun		Adj	
Party	Dem	Rep	Dem	Rep
1900	5.943880e-07	6.111457e-07	0.000020	0.000039
1901	6.267058e-07	6.308210e-07	0.000021	0.000039
1902	5.926260e-07	5.927591e-07	0.000020	0.000040
1903	5.866815e-07	5.819056e-07	0.000020	0.000039
1904	5.369485e-07	5.179634e-07	0.000019	0.000036

# groupby, agg를 활용한 'The Democrats - The Republicans', 'Democratic - Republican' 비교

```
dem_rep_sub = US_dem_rep_NOUN_ADJ.groupby(level='Class', axis=1).agg(lambda x: x.iloc[:,0]-x.iloc[:,1
dem_rep_sub.index.name = 'YEAR'
dem_rep_sub.plot()
plt.axis([1940,2000,-0.5e-5, 0.5e-5])
plt.hlines(0, 1900, 2000, linestyle='-.', color='grey')
plt.title('NOUN & ADJ', weight='bold')
```

Text(0.5, 1.0, 'NOUN & ADJ')

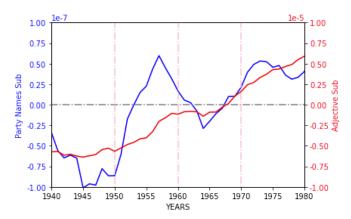


두 차트는 비슷한 변화양상을 보이는 듯하나, 포착되는 빈도에서 막대한 차이를 보인다. 이를 통해 미국사회의 정치적 의제 변화를 반영하는 형용사형이 정당명을 나타내는 명사형보다 더 풍부한 데이터로 이루어져 있음을 확인할 수 있다. 다음으로는 두 그래프를 서로 다른 y축에서 비교하여 상이한 데이터의 추세양상을 확인하였다.

# subplots를 활용한 'The Democrats - The Republicans', 'Democratic - Republican' 비교

```
a = US_dem_rep_PARTY['DEM'] - US_dem_rep_PARTY['REP']
b = US_dem_rep_ADJ['DEM'] - US_dem_rep_ADJ['REP']
fig, ax1 = plt.subplots()
ax1.set xlabel('YEARS')
ax1.set_ylabel('Party Names Sub', color='blue')
ax1.set_xlim([1940,1980])
ax1.set_ylim([-1e-7,1e-7])
ax1.plot(a, color='blue')
ax1.tick_params(axis='y', labelcolor='blue')
ax2 = ax1.twinx()
ax2.set_ylabel('Adjective Sub', color='red')
ax2.set_ylim([-1e-5,1e-5])
ax2.plot(b, color='red')
ax2.tick_params(axis='y', labelcolor='red')
plt.hlines(0, 1900, 2000, linestyle='-.', color='grey')
plt.vlines(1950, -1, 1, linestyle='-.', color='pink')
plt.vlines(1960, -1, 1, linestyle='-.', color='pink')
plt.vlines(1970, -1, 1, linestyle='-.', color='pink')
```

<matplotlib.collections.LineCollection at 0x7fb445b92ee0>



두 데이터의 라인 차트는 서로 유사한 모습으로 나타난다. 단, 정당 이름을 나타내는 명사형은 미국 내정치적 사건에 따라 극단적으로 변동하는 반면, 형용사형은 비교적 일관되고 안정적인 모습을 띤다. 이는 대중의 정치적 담론 변화가 양당에 대한 명성 변화보다 점진적으로 이루어지는 현실을 반영한다. 할리우드 10인에 대한 억압이 지속된 1950~1960년 구간을 확인했을 때, 명사형 차트(파랑)로는 유의미한 특성을 발견하기 어렵다. 하지만 형용사형 차트(빨강)는 해당 기간에 메카시즘을 주도한 공화당의 영향력이 우세했으며, 동시에 개인의 권리보다 공동체의 이익을 우선시하는 공화주의적 정치의제가 주목받았음을 보여준다. 할리우드 10인에 대한 명성 회복이 이루어진 1960~1970년 기간의 경우, 명사형차트(파랑)는 급격한 하락 및 상승 양상을 보여 할리우드 10인의 명성과 연관성을 발견하기 어렵다. 그러나 형용사형 차트(빨강)는 일관되게 0 근처의 값을 유지하다가 상승하는 양상을 보인다. 이는 미국사회에서 민주주의적 정치의제에 대한 관심이 공화주의와 동일한 정도에 도달했음을 보여주며, 나아가기존의 공화주의적 담론과 정책에 대한 새로운 평가가 이루어질 수 있음을 암시한다.

## 5. 결론

앞서 살펴본 'Democratic - Republican' 차트는, 정치적 억압집단의 명성 회복 시기가 가해 정당의 정치사상과 대립 정당의 정치사상에 대한 사회적 관심도로 결정된다는 사실을 보여준다. 필자는 억압 주체가 대변하는 정치사상(형용사)의 영향력이 압도적으로 높을 경우, 억압 집단에 대한 명성회복인 이루어지기 어렵다고 판단했다. 또한 그와 대립되는 정치사상의 영향력이 기존 주류사상과 대등한 정도에 도달할 경우, 억압집단에 대한 재평가와 명성 회복이 이루어질 수 있음을 확인하였다. 이러한 결론은 추후 양당체제 국가의 정치적 억압사례를 추가적으로 조사하는 일반화 과정이 필요하며, Democratic 및 Republican 이외의 다양한 형용사가 대변하는 사회적 의제에 대해서도 추가적인 확인이 이루어져야 한다.