

コーパスから見える
20世紀韓国の言葉・社会

南潤珍
東京外国語大学准教授


1. 韓国のコーパス事情

- 21世紀世宗計画
世宗コーパス
- 延世大学言語情報研究院
延世20世紀韓国語コーパス
- 高麗大学
SJ-RIKS 拡張版
高麗話し言葉コーパス

21世紀世宗計画: 世宗コーパス

- 1998年～2007年
- 韓国政府(文化体育観光部)
- 構成:
 - 書き言葉コーパス: 36,879,143語
 - 話し言葉コーパス: 805,646語
 - パラレル・コーパス
 - 歴史コーパス


<https://ithub.korean.go.kr/user/main.do>



延世20世紀韓国語コーパス

- 延世大学 1989年～辞書編纂のため
- 1900年代～1990年代10年単位
- 新聞・文学・情報・辞書・雑誌・教科書
- 1億5千万語


<https://ilis.yonsei.ac.kr/corpus/>



世宗民研コーパス(拡張版)

- 世宗コーパスの修正・補完
- 1億1600万語

<http://riksdb.korea.ac.kr/>



高麗話し言葉コーパス

- 1926年～ 1942年:58,756語
- 1948年～1999年:457,913語
- 映画・ドラマ・漫画・普通学校読本など
- 未公開

検索してみると…

- 20世紀韓国語コーパス
<https://ilis.yonsei.ac.kr/corpus/20century>
- 世宗民研コーパス
<http://riksdb.korea.ac.kr/>
- 世宗コーパス
<https://ithub.korean.go.kr/user/main.do>

何のためのコーパスなのか？

- 仮説や主張の根拠として：出所が分かる資料
- 頻度調査ができる：出現回数そのものも大事であるが，その頻度の意味が解釈できなければ全体で占める示す割合など。

2. 頻度を利用した研究例(1)

- 日本語の「やさしい」に対応する韓国語

1. 부드럽다
2. 착하다
3. 친절하다
4. 다정하다
5. 자상하다
6. 상냥하다

- 世宗コーパスから共起語を調査:コロケーション
- 各単語の後ろに来る名詞を調べた

1. 「부드러운」의 코로케이션: 前項語彙(+1位置)

| | | | |
|---------|----|----------|----|
| 목소리로 声で | 43 | 바람이 | 10 |
| 느낌을 感じ | 22 | 것, | 9 |
| 맛을 味 | 22 | 말로 言葉で | 9 |
| 미소를 微笑 | 17 | 꼭선으로 | 8 |
| 것이 | 16 | 꼭선을 | 8 |
| 분위기를 | 15 | 그녀의 | 7 |
| 목소리를 声を | 14 | 손길로 | 7 |
| 어조로 語調で | 14 | 음성으로 音声で | 7 |
| 것을 | 10 | 느낌의 感じの | 6 |

2. 「착한」의 코로케이션: 前項語彙(+1位置)

| | | | |
|----------|----|-----------|----|
| 일을 | 89 | 사람들이 人たちが | 12 |
| 사람이 人が | 44 | 심성을 | 9 |
| 사람은 人は | 23 | 사람에게 人に | 8 |
| 사람을 人を | 22 | 사람의 人の | 8 |
| 마음을 | 19 | 여자 女 | 8 |
| 사마리아인 | 18 | 것 | 7 |
| 일 | 17 | 마음씨가 | 7 |
| 아이기가 子供が | 15 | 마음씨를 | 7 |
| 사람 | 13 | 아이를 子供を | 7 |
| 어린이가 子供が | 13 | 아기 赤ちゃん | 6 |

3. 「친절한」의 코로케이션: 前項語彙(+1位置)

| | | | |
|---------|---|----------|---|
| 설명이 説明が | 6 | 서비스가 | 3 |
| 태도를 態度を | 6 | 안내서가 | 3 |
| 말 | 5 | 가이드를 | 2 |
| 설명에 説明に | 5 | 공무원에게는 | 2 |
| 설명을 説明を | 5 | 노파는 | 2 |
| 안내를 案内を | 5 | 대접을 | 2 |
| 태도로 態度で | 5 | 마음씨 | 2 |
| 설명도 説明も | 4 | 마음으로 | 2 |
| 표정을 表情を | 4 | 말씨로 話し方で | 2 |
| 마음이 | 3 | 말씨를 話し方を | 2 |

4. 「다정한」의 코로케이션: 前項語彙(+1位置)

| | | | |
|---------|----|----------------|---|
| 모습을 | 12 | 얼굴로 | 4 |
| 목소리로 | 9 | 연인처럼恋人のよ うに | 4 |
| 친구 友達 | 9 | 우리 | 4 |
| 친구가 友達が | 7 | 이웃이 隣人が | 4 |
| 모습으로 | 6 | 느낌을 | 3 |
| 미소를 | 5 | 대화를 対話を | 3 |
| 눈길을 | 4 | 모습 | 3 |
| 말 | 4 | 웃음을 | 3 |
| 말을 | 4 | 친구들과 友達と | 3 |
| 모습이 | 4 | 포즈를 | 3 |

5. 「자상한」의 코로케이션

| | | | |
|----------------|---|-----------------|---|
| 마음 | 3 | 부모의 | 2 |
| 설명을 | 3 | 분이었다. 方 だった | 2 |
| 성격의 | 3 | 사랑을 | 2 |
| 아버지 お父さん | 3 | 선생님은 先 生は | 2 |
| 아버지의 お父さ んの | 3 | 손윗사람으로 目上の人で | 2 |
| 남편, 夫 | 2 | 아버지, お父 さん | 2 |
| 남편이 夫が | 2 | 아버지와 お 父さんと | 2 |
| 마음씀일 | 2 | 안내에 | 2 |
| 면이 | 2 | 어머니가 お 母さんが | 2 |
| 목소리로 | 2 | 해설을 | 2 |

6. 「상냥한」의 코로케이션: 前項語彙(+1位置)

| | | | |
|------|---|------------------|---|
| 목소리로 | 8 | 아가씨의 女性の | 1 |
| 미소를 | 5 | 아나운서가 アナウンサーが | 1 |
| 목소리의 | 3 | 아씨만 お嬢様だ け | 1 |
| 표정을 | 3 | 안내양, 案内嬢 | 1 |
| 눈초리로 | 2 | 그 | 1 |
| 목소리는 | 2 | 그녀는 彼女は | 1 |
| 어조로 | 2 | 낮이 | 1 |
| 여자 女 | 2 | 내 | 1 |
| 웃음을 | 2 | 눈빛을 | 1 |
| 친절을 | 2 | 눈이 | 1 |

韓国語 「優しい」 語群の用法の分類

| | 対象 | 関係性 | 人間(男女) | 年齢 |
|---------|----|-----|--------|----|
| 1. 부드러운 | | 無 | - | - |
| 2. 착한 | 無 | 無 | 人間(女性) | 下 |
| 3. 친절한 | 有 | 無 | - | - |
| 4. 다정한 | 無 | 有 | - | - |
| 5. 자상한 | 有 | 有 | 人間(男性) | 上 |
| 6. 상냥한 | 無 | 無 | 人間(女性) | - |

• 카요는 귀엽고 6상냥한 여자친구이며, 마사시는 작은 것에도 신경 써 주는 5자상한 남자친구이다.<동아일보 2005.2.7>

佳代は可愛くて優しい彼女であり、正志は些細なことにも配慮してくれる優しい彼した。

• 동생도 5형의 자상한 배려로 교육훈련에 열심히 임하고 있다<동아일보 2008.2.4>
弟も兄の優しい配慮のもと、教育訓練に熱心に臨んでいる

• 김양은 엄마 말 잘 듣는 2착한 아이였다<동아일보 2008.7.28>
金さんはお母さんのいうことを良く聞くよい子だった。

| | 対象 | 関係性 | 人間(男女) | 年齢 |
|---------|----|-----|--------|----|
| 1. 부드러운 | | 無 | - | - |
| 2. 착한 | 無 | 無 | 人間(女性) | 下 |
| 3. 친절한 | 有 | 無 | - | - |
| 4. 다정한 | 無 | 有 | - | - |
| 5. 자상한 | 有 | 有 | 人間(男性) | 上 |
| 6. 상냥한 | 無 | 無 | 人間(女性) | - |

3. 頻度を利用した研究例(2)

- 韓国語の一般否定文
 - 否定述語(않다)によるもの(1)
 - 否定副詞(안)によるもの(2)
 - 2つの否定文はどのような使用頻度で現れるのか?
- 韓国語の能力否定文
 - 否定述語(못하다)によるもの(1)
 - 否定副詞(못)によるもの(2)
 - 第3の否定(ㄹ 수 없다)(3)

世宗コーパスから抽出した小説テキスト
サンプルコーパスの内訳

| 年代 | 作品種数 | 大きさ(語) | 年代 | 作品種数 | 大きさ(語) |
|--------|------|---------|--------|------|-----------|
| 1900年代 | 13 | 200,070 | 1950年代 | 39 | 197,350 |
| 1910年代 | 14 | 193,534 | 1960年代 | 31 | 200,259 |
| 1920年代 | 51 | 200,108 | 1970年代 | 28 | 200,017 |
| 1930年代 | 42 | 252,555 | 1980年代 | 30 | 200,067 |
| 1940年代 | 30 | 147,905 | 1990年代 | 37 | 200,100 |
| 合計(全体) | | | | 315 | 1,991,965 |

3.1. 否定述語의 形態

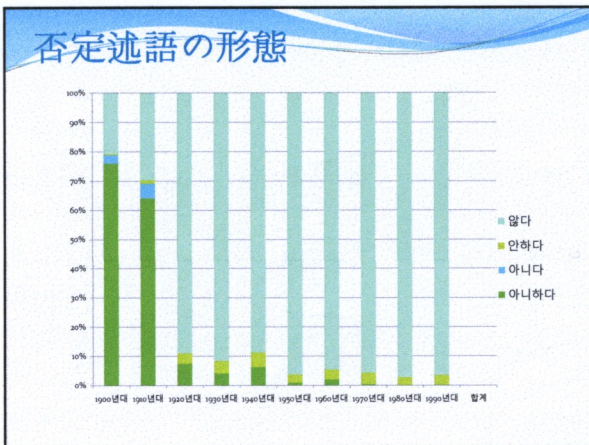
- 否定述語の変化
‘아니하-’ > ‘않-’ : 19C

20世紀初の形態

- 1) 아니하-
- 2) 아니-
- 3) 안하-
- 4) 않-

否定叙述語의 形態

| 形態 年代 | 1)아니하다 | 2)아니다 | 3)안하다 | 4)않다 | 合計 |
|----------|--------|-------|-------|------|------|
| 1900 | 948 | 37 | 3 | 260 | 1248 |
| 1910 | 729 | 58 | 13 | 339 | 1139 |
| 1920 | 113 | 0 | 53 | 1352 | 1518 |
| 1930 | 87 | 0 | 91 | 1966 | 2144 |
| 1940 | 75 | 0 | 58 | 1060 | 1193 |
| 1950 | 16 | 0 | 52 | 1770 | 1838 |
| 1960 | 36 | 0 | 56 | 1663 | 1755 |
| 1970 | 5 | 0 | 57 | 1425 | 1487 |
| 1980 | 3 | 0 | 44 | 1686 | 1733 |
| 1990 | 3 | 0 | 61 | 1805 | 1869 |



- 1)아니하-의 頻度는 1920年代に入つて急激に減少.
- 同時に 4)않-의 頻度が急増
- 新小説と現代小説の文体的違ひが否定叙述語の携帶にも反映される
- 19世紀に再構造化された '않다' が1920年代に入つて否定叙述語としての安定した地位を獲得していく.

3.2. 韓国語の一般否定文

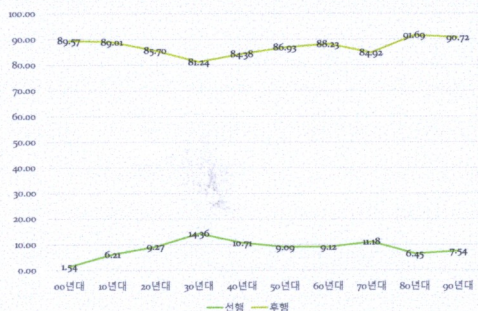
- 否定述語(않다)によるもの(1)
- 否定副詞(안)によるもの(2)
- 2つの否定文はどのような使用頻度で現れるのか?

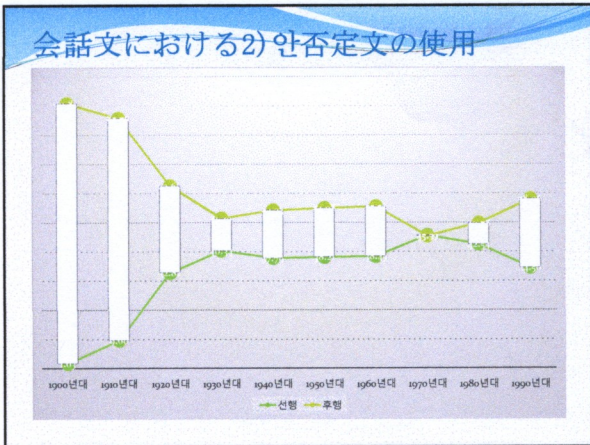
テキストの内部構造と否定文の使用

テキスト類型と否定文の類型別使用頻度

| | 会話文 | | | | | | 地の文 | | | | | |
|------|--------|-------|--------|-------|-------|------|--------|-------|--------|-------|-------|------|
| | 2 先行否定 | | 1 後行否定 | | X時点否定 | | 2 先行否定 | | 1 後行否定 | | X時点否定 | |
| | 頻度 | 比率 | 頻度 | 比率 | 頻度 | 比率 | 頻度 | 比率 | 頻度 | 比率 | 頻度 | 比率 |
| 1900 | 11 | 1.63 | 612 | 90.80 | 51 | 7.57 | 11 | 1.63 | 612 | 90.80 | 51 | 7.57 |
| 1910 | 64 | 9.60 | 572 | 85.76 | 31 | 4.65 | 64 | 9.60 | 572 | 85.76 | 31 | 4.65 |
| 1920 | 172 | 32.82 | 328 | 62.60 | 24 | 4.58 | 172 | 32.82 | 328 | 62.60 | 24 | 4.58 |
| 1930 | 259 | 40.34 | 330 | 51.40 | 53 | 8.26 | 259 | 40.34 | 330 | 51.40 | 53 | 8.26 |
| 1940 | 161 | 37.79 | 231 | 54.23 | 34 | 7.98 | 161 | 37.79 | 231 | 54.23 | 34 | 7.98 |
| 1950 | 230 | 38.21 | 331 | 54.98 | 41 | 6.81 | 230 | 38.21 | 331 | 54.98 | 41 | 6.81 |
| 1960 | 254 | 38.43 | 367 | 55.52 | 40 | 6.05 | 254 | 38.43 | 367 | 55.52 | 40 | 6.05 |
| 1970 | 253 | 45.59 | 252 | 45.41 | 50 | 9.01 | 253 | 45.59 | 252 | 45.41 | 50 | 9.01 |
| 1980 | 194 | 42.54 | 227 | 49.78 | 35 | 7.68 | 194 | 42.54 | 227 | 49.78 | 35 | 7.68 |
| 1990 | 167 | 34.50 | 281 | 58.06 | 36 | 7.44 | 167 | 34.50 | 281 | 58.06 | 36 | 7.44 |

地の文における안否定文の使用



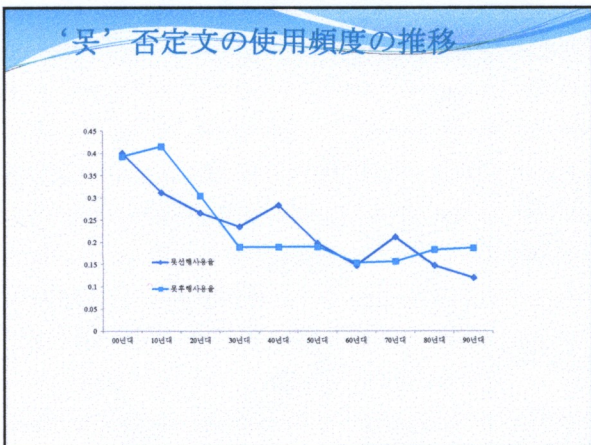


- 회話文と地の文ともに1)の使用率が高い
- 회話文では1)と2)頻度差が地の文でのそれより顕著に減少
- 1920年代から会話文での2)の使用が急激に増加
- 1970年代, 1980年代にも2)の使用が増加

- 3.3. 韓国語の能力否定文
- 否定述語(못하다)によるもの(1)
 - 否定副詞(못)によるもの(2)
 - 第3の否定(ㄹ 수 없다)(3)

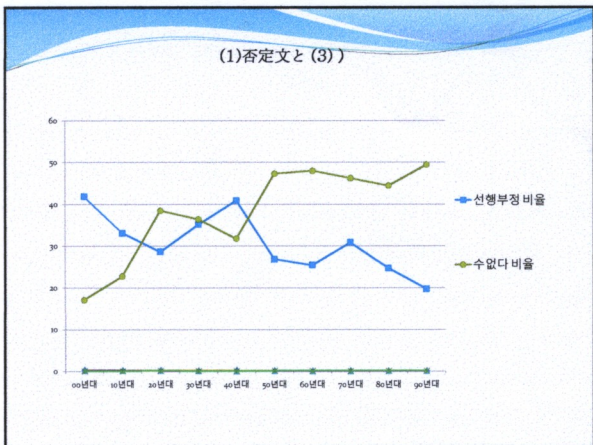
‘ㄹ’ 否定文の使用頻度の推移

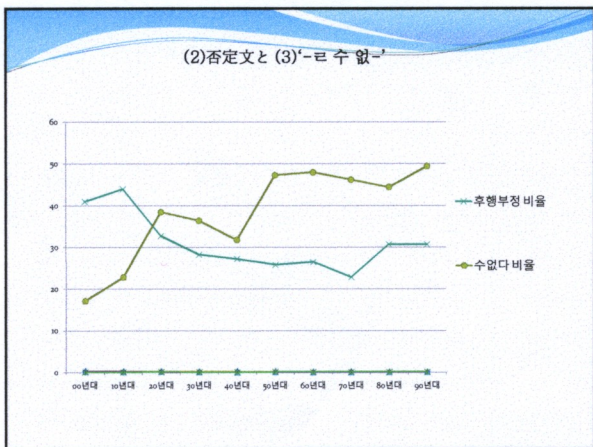
| | (2) 先行否定 | | (1) 後行否定 | | | (2) 先行否定 | | (1) 後行否定 | |
|------|----------|------|----------|-------|------|----------|------|----------|------|
| | 頻度 | 比率 | 頻度 | 比率 | | 頻度 | 比率 | 頻度 | 比率 |
| 1900 | 802 | 0.41 | 787 | 0.393 | 1950 | 397 | 0.20 | 381 | 0.19 |
| 1910 | 627 | 0.31 | 833 | 0.42 | 1960 | 296 | 0.15 | 308 | 0.15 |
| 1920 | 533 | 0.27 | 610 | 0.30 | 1970 | 425 | 0.21 | 314 | 0.16 |
| 1930 | 471 | 0.24 | 379 | 0.19 | 1980 | 295 | 0.15 | 367 | 0.18 |
| 1940 | 569 | 0.28 | 380 | 0.19 | 1990 | 241 | 0.12 | 375 | 0.19 |



(1)(2) 否定文と (3)

| | (2) 先行否定 | | (1) 後行否定 | | (3) ㄹ 수 없다 | |
|------|----------|-------|----------|-------|------------|-------|
| | 頻度 | 比率 | 頻度 | 比率 | 頻度 | 比率 |
| 1900 | 802 | 41.84 | 786 | 41.00 | 329 | 17.16 |
| 1910 | 627 | 33.16 | 833 | 44.05 | 431 | 22.79 |
| 1920 | 533 | 28.67 | 610 | 32.81 | 716 | 38.52 |
| 1930 | 471 | 35.20 | 379 | 28.33 | 488 | 36.47 |
| 1940 | 569 | 40.88 | 380 | 27.30 | 443 | 31.82 |
| 1950 | 397 | 26.88 | 381 | 25.80 | 699 | 47.33 |
| 1960 | 296 | 25.47 | 308 | 26.51 | 558 | 48.02 |
| 1970 | 425 | 30.91 | 314 | 22.84 | 636 | 46.25 |
| 1980 | 295 | 24.75 | 367 | 30.79 | 530 | 44.46 |
| 1990 | 241 | 19.77 | 375 | 30.76 | 603 | 49.47 |






- (3)의 頻度는 1920年代, 1950年代 2回에 增加
- 1920年代는 一般否定文においても 能力否定文においても 変化を見せる時期: 何かあったの?
- 1950年代に 能力否定の(3)タイプが 増加: 何かあったのか?

- 1920年代;小説における文体の変化が起こり,現代小説が定着
- 1950年代;1945年以後アメリカ文化が拡散
1945年~1950年の出版動向-翻訳本の出版が急増
英語翻訳体の拡散
 - 能力否定の(1)(2)は形容詞述語では作れない。
 - 能力否定の(3)は形容詞述語で作れる
더 이상 좋을 수는 없다
「It couldn't be better」「この上なく良い」

4. 終わりに

韓国のコーパス事情
コーパスで何を見つけるのか?
頻度と言葉の研究
頻度から見えてくること
言葉だけを見ては言葉が見えない

 ご清聴ありがとうございました!
