

2026 코퍼스언어학 하계 워크숍 (2026.08.28)

# 코퍼스와 AI의 융합 연구

권혁승

# 2026 코퍼스 언어학 하계 워크숍



- 대상: 교사, 예비교사, 교수, 강사, 연구원, 대학(원)생, 기타 등
- 주최: 아시아태평양언어문화교육학회, 한국중등영어교육학회
- 주관: 국립인천대학교 코퍼스연구소
- 협력: 서울대학교
- 일시: 2026년 8월 28일(금) 오후 1시 - 5시
- 장소: 서울대학교
- 등록: <https://forms.gle/WdGB1AnMBu8Th6C97>
- 등록 마감: 2026년 8월 14일(금) 오후 6시
- 기타: 8월 21일(금) 오전 9시에 사전 등록 신청자에게 문자 발송 예정. 장소 안내 문자를 받지 못하신 경우, 이메일로 문의해 주시기 바랍니다. 본 워크숍은 온라인 실시간 중계 없이 오프라인으로만 진행합니다.
- 안내: <https://apalce.org/conferences>
- 문의: [office@apalce.org](mailto:office@apalce.org)

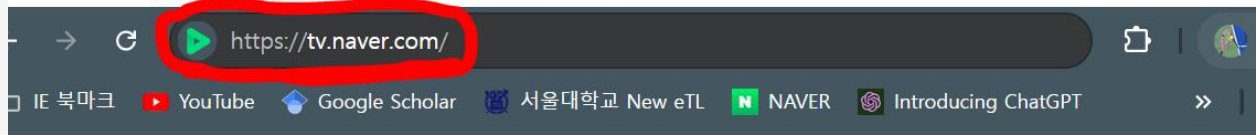
| 시간            | 프로그램                                |
|---------------|-------------------------------------|
| 13:00 - 13:10 | 개회                                  |
| 13:10 - 14:40 | 코퍼스와 AI의 융합 연구<br>권혁승 교수<br>(서울대학교) |
| 14:40 - 14:50 | 휴식                                  |
| 14:50 - 16:30 | AntConc 4.4<br>정채관 교수<br>(국립인천대학교)  |
| 16:30 - 17:00 | 종합토론 및 질의응답                         |

주최  
 아시아태평양언어문화교육학회  
 Asia-Pacific Association of Language, Culture, and Education  
 KASEE THE KOREA ASSOCIATION OF  
 SECONDARY ENGLISH EDUCATION

주관  
 INU 국립인천대학교 코퍼스연구소  
 Institute for Corpus Research, Incheon National University

협력  
 서울대학교  
 Seoul National University

# 코퍼스언어학: NAVER TV 강의 (선수 강의)



NAVER TV

코퍼스언어학

취소



## Q2. '코퍼스 언어학'이란?

언어학의 영역은 언어를 구성하는 언어 구성 ...

#서울대지식교양 #생각의열쇠 #천개의키워드

2년 전 · ▷ 870



서울대 지식교양 강연 - 생각의 열쇠



## 코퍼스언어학 - 서울대 영어영문학과 권혁승.

20세기 후반에 컴퓨터와 함께 급속히 발달한 ...

#서울대지식교양 #생각의열쇠 #천개의키워드

2년 전 · ▷ 2,439



서울대 지식교양 강연 - 생각의 열쇠

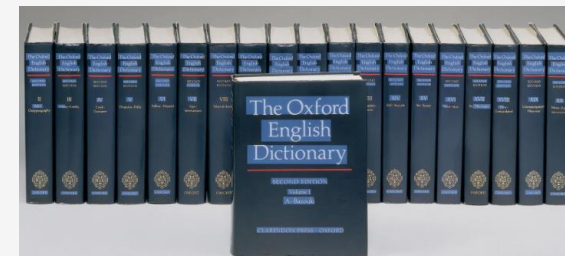
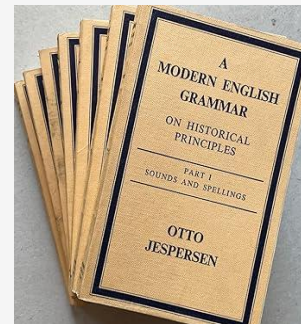
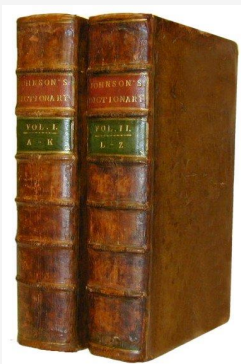
## 강연목차

01. '코퍼스'란 무엇인가?
02. '코퍼스 언어학'이란 무엇인가?
03. 코퍼스 구축의 발달과정에 대해
04. 코퍼스의 유형
05. 코퍼스 분석의 기본개념 3가지
06. '콘코던스(Concordance)'란?
07. '콜로케이션(Collocation)'이란?
08. 코퍼스 분석이 '사전편찬'에 어떻게 사용되는가?
09. 코퍼스 분석이 '문법 기술'에 어떻게 사용되는가?
10. 키워드 분석이란 무엇이며 어떻게 활용되는지?

# 목차

---

- 코퍼스의 기원
- 코퍼스의 발달 과정
- 코퍼스 분석의 기본 개념
- 온라인 코퍼스와 AI 도입
- 코퍼스와 AI



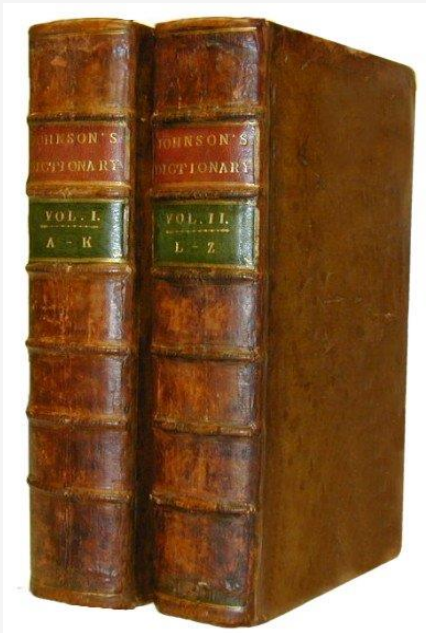
AI/LLM insights integrated into corpus

Using Corpora and AI for Second Language Research

Implementing AI Tools  
for Language Teaching  
and Learning

# 코퍼스의 기원: 사전 편찬

- Johnson's *Dictionary of the English Language* (1755)

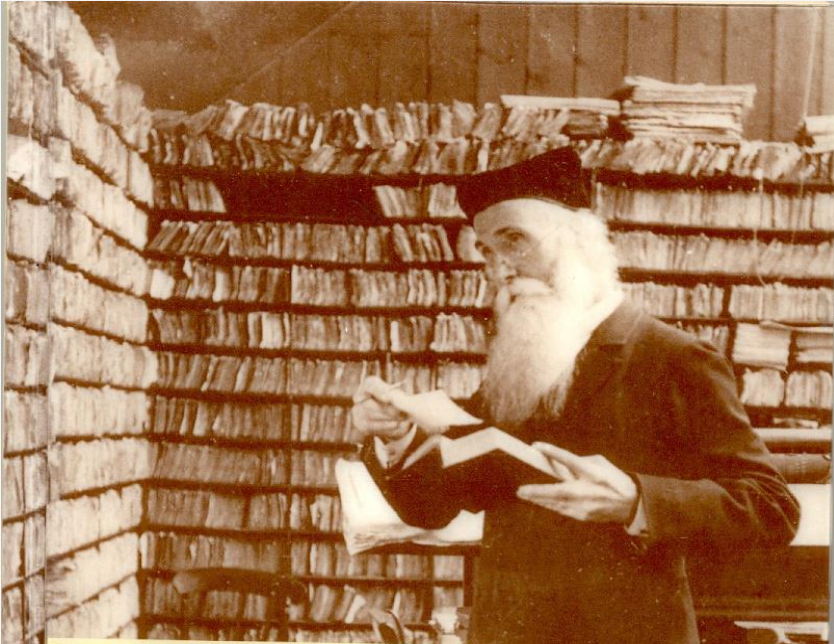


OATS. *n. f.* [*aēn*, Saxon.] A grain, which in England is generally given to horses, but in Scotland supports the people.  
It is of the grass leaved tribe ; the flowers have no petals, and are disposed in a loose panicle : the grain is eatable.  
The meal makes tolerable good bread. *Miller.*  
The *oats* have eaten the horses. *Shakespeare.*  
It is bare mechanisim, ~~no~~ otherwise produced than the turning of a wild *oatbeard*, by the insinuation of the particles of moisture. *Locke.*  
For your lean cattle, fodder them with barley straw first, and the *oat* straw last. *Mortimer's Husbandry.*  
His horse's allowance of *oats* and beans, was greater than the journey required. *Swift.*

## 코퍼스의 기원: 사전 편찬

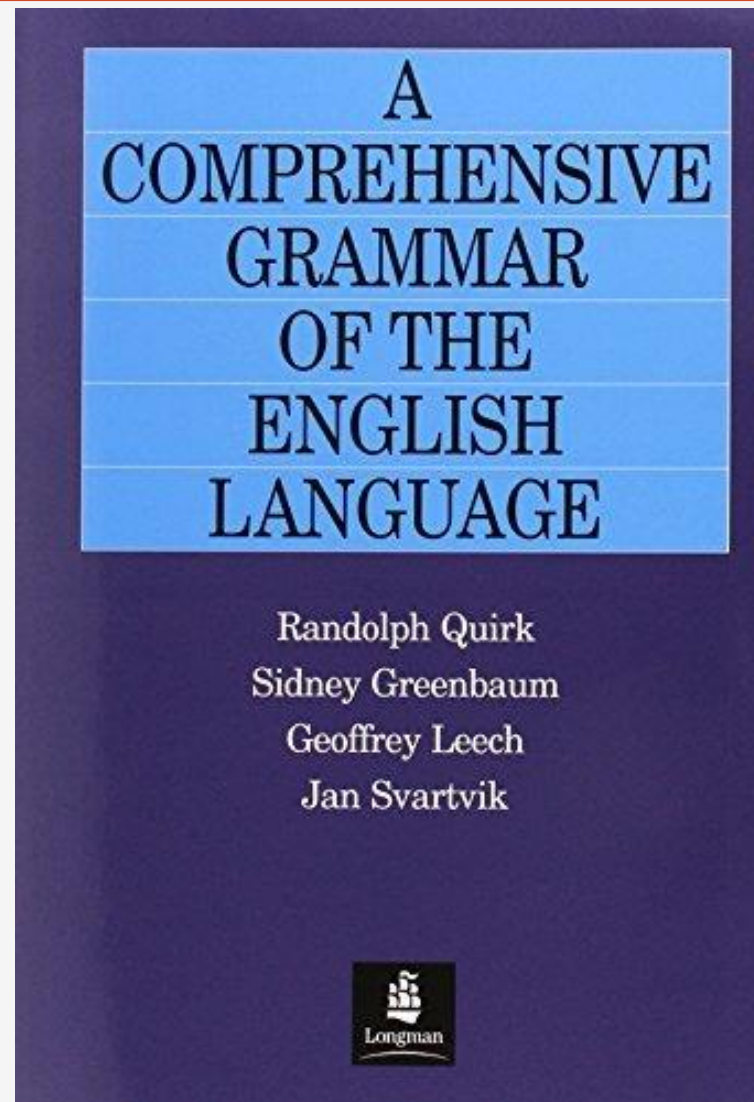
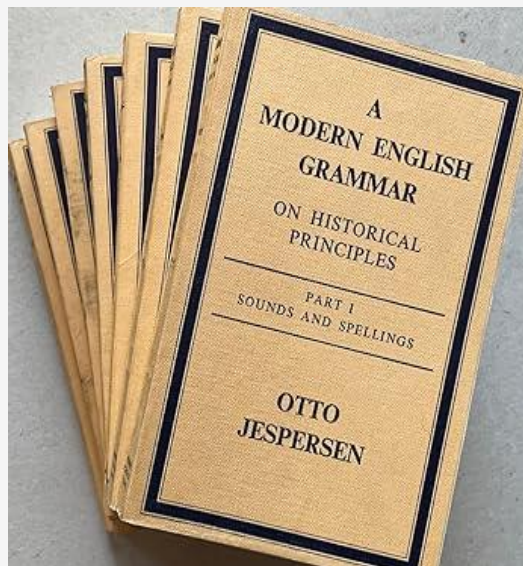
---

- Murray's *Oxford English Dictionary* (1878-1928)



## 코퍼스의 기원: 문법서 편찬

- Jespersen's *A Modern English Grammar on Historical Principles*



ist and grammarian, best  
Grammar on Historical  
sive, manual, and empirical  
sands of shoeboxes with slips  
conversation. [Wikipedia +4](#)

e:  
er-based corpus linguistics,  
f English usage by reading  
natical structures on slips of

slips were organized into  
hagen.

dbreaking because it relied  
) rather than just theoretical

ned the foundation for his  
ge, including *A Modern*

ResearchGate +5

# 코퍼스(언어학)의 발달 과정

- First-generation corpus (1960s and 1970s)
  - Brown Corpus, LOB Corpus (1 million)
- Second-generation corpus (1980s)
  - Birmingham Collection of English Text (18 million)
- Third-generation corpus (1990s)
  - The British National Corpus (100 million)
  - The Bank of English (over 500 million)
- Next generation corpus (2000s & 2010s)
  - One million million word corpora before 2021
- Current online corpora (2020s)



English-Corpora.org

corpora PDF guides videos related resources users my account upgrade help

Overview: brief | detailed

Now available: AI/LLM insights integrated into corpus results

These are the most widely used online corpora, and they serve many different purposes for teachers and researchers at universities throughout the world. In addition, the corpus data (e.g. full-text, word frequency) has been employed by a wide range of companies in many different fields, especially technology and language learning.

The links below are for the free online interface. You can also purchase and download the corpora for use on your own computer.

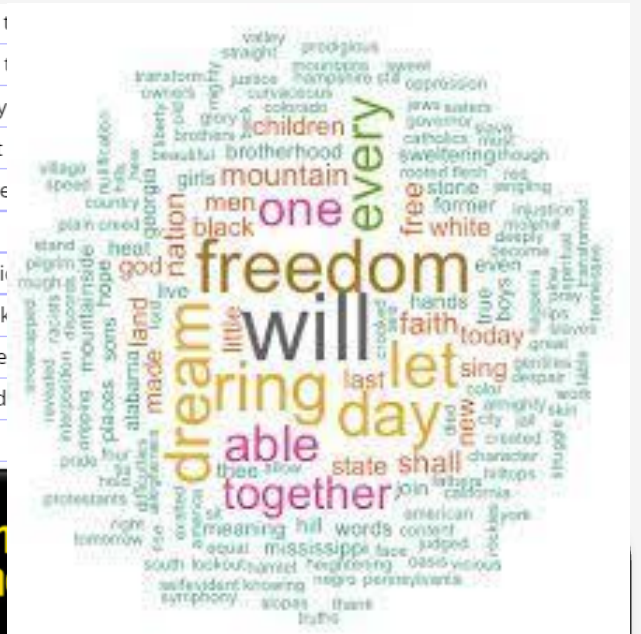
| Corpus                                         | Download | # words       | Dialect      | Time period    | Genre(s)         |
|------------------------------------------------|----------|---------------|--------------|----------------|------------------|
| News on the Web (NOW)                          | Download | 24.4 billion+ | 20 countries | 2010-yesterday | Web: News        |
| iWeb: The Intelligent Web-based Corpus         | Download | 14 billion    | 6 countries  | 2017           | Web              |
| Global Web-Based English (GloWbE)              | Download | 1.9 billion   | 20 countries | 2012-2013      | Web (incl blogs) |
| Wikipedia Corpus                               | Download | 1.9 billion   | (Various)    | 2014           | Wikipedia        |
| Coronavirus Corpus                             | Download | 1.5 billion   | 20 countries | 2020-2023      | Web: News        |
| Corpus of Contemporary American English (COCA) | Download | 1.0 billion   | American     | 1990-2019      | Balanced         |
| Corpus of Historical American English (COHA)   | Download | 475 million   | American     | 1820-2019      | Balanced         |
| The TV Corpus                                  | Download | 325 million   | 6 countries  | 1950-2018      | TV shows         |
| The Movie Corpus                               | Download | 200 million   | 6 countries  | 1930-2018      | Movies           |
| Corpus of American Soap Operas                 | Download | 100 million   | American     | 2001-2012      | TV shows         |
|                                                |          |               |              |                |                  |
| Hansard Corpus                                 |          | 1.6 billion   | British      | 1803-2005      | Parliament       |
| Early English Books Online (EEBO)              |          | 755 million   | British      | 1470s-1690s    | (Various)        |
| Corpus of US Supreme Court Opinions            |          | 130 million   | American     | 1790s-2017     | Legal opinions   |
| TIME Magazine Corpus                           |          | 100 million   | American     | 1923-2006      | Magazine         |
| British National Corpus (BNC) *                |          | 100 million   | British      | 1980s-1993     | Balanced         |
| Strathy Corpus (Canada)                        |          | 50 million    | Canadian     | 1970s-2000s    | Balanced         |
| CORE Corpus                                    |          | 50 million    | 6 countries  | 2014           | Web              |
| From Google Books n-grams (compare)            |          |               |              |                |                  |
| American English                               |          | 155 billion   | American     | 1500s-2000s    | (Various)        |
| British English                                |          | 34 billion    | British      | 1500s-2000     | (Various)        |

# 코퍼스 분석의 기본 개념

- 가장 중요한 세 가지 기본 개념:
  - 어휘 빈도수(word frequency)
  - 콘코던스(concordance)
  - 콜로케이션(collocation)
- 그 외에도
  - 어휘목록(wordlist)
  - 키워드(keyword)
  - n-그램(n-gram)
  - 클러스터(cluster)
  - 플롯(plot)

his. I believe we need to take a break from each other. We both need  
this alone ... But if you'll take a chance on me, we can do great things  
?; and much more. We take a closer look at the increase in attacks by  
n concrete. Or we could take a fresh look at our challenges and adjust  
out a father, nobody will take a girl. I had to find my husband and have  
ould make a difference. Take a left here. "A string of cars in front of  
about that on his show. Take a look. BILL-O'REILLY# We have rock soli  
we're going to have you take a look. Let's start with the must-see look  
When you point a finger, take a look at your hand because there's three  
g ... even at night. Let's take a look at the golf case. Positive for blood  
ctor's secretary? Jeeze, take a minute to think. " # " Wait a second.  
, keep it on. Let me just take a picture. [speaking]  
ndary Uncle Drew, they take a road trip to  
n goes into rehab. Let's take a short break  
d over," said Jamaal. # " Take a shower. " I  
you need to go (m) (m) Take a step back (m)  
ner. He watched Mandy take a step away from  
rant. So break a finger, take a tooth. do v  
. # " Anyone who would take a woman like  
ction plus a firewall and take advantage of

notes  
a test  
control  
medicine  
a shower  
someone's place  
a chance  
actions  
a bus  
measures  
a sip  
T  
a test  
control  
medicine  
a shower  
someone's place  
a chance  
actions  
a bus  
measures  
a sip  
someone's advice  
someone's temperature  
a deep breath  
responsibility  
offence  
a train  
office  
power  
a seat



# 단어 빈도수(word frequency)

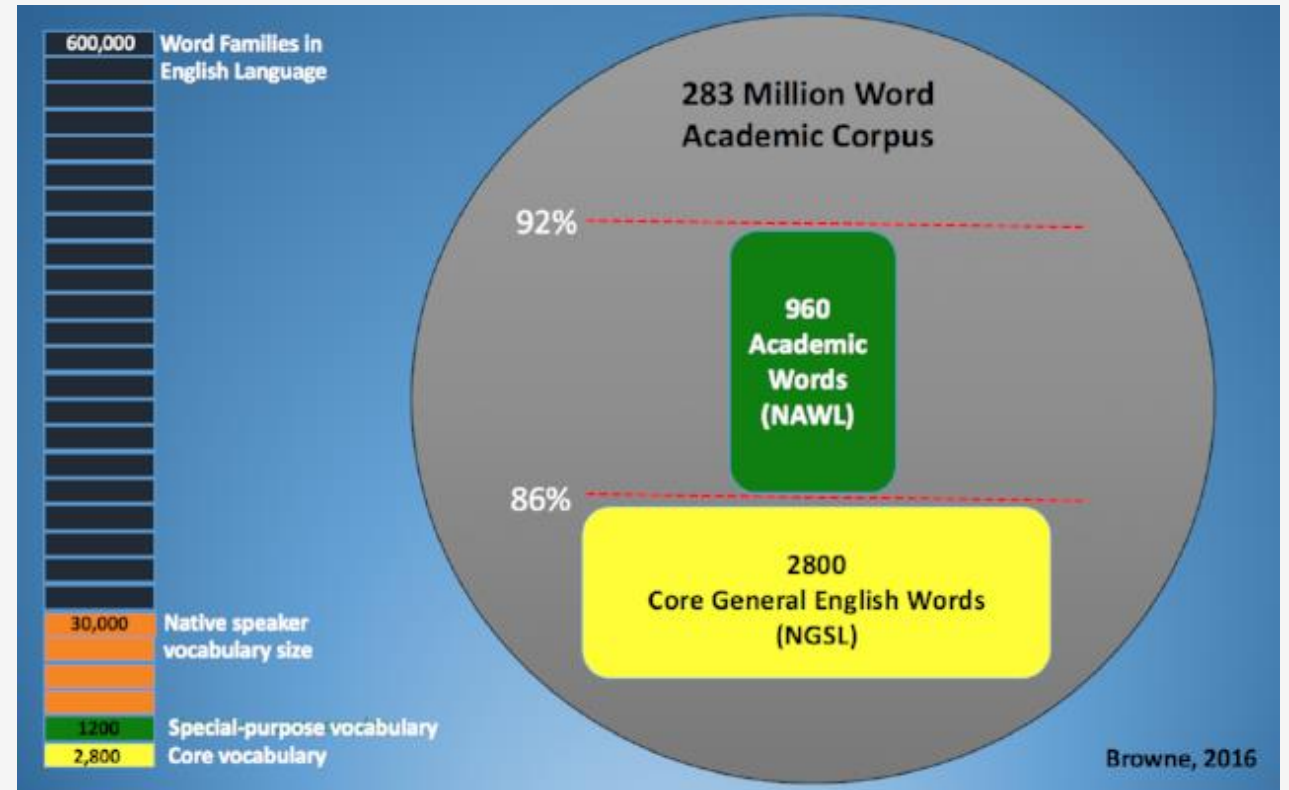
- 단어 빈도수: 코퍼스에서 단어의 출현 횟수
- 영어에서 가장 많이 사용되는 단어는 무엇일까?
- 상위 고빈도 단어의 비중(percentage)은 어떻게 될까?
- Brown Corpus를 예로 들어 생각해 봅시다.

## Brown 코퍼스의 최상위 빈도수 단어

|    |      |       |        |       |
|----|------|-------|--------|-------|
| 1  | the  | 69454 | (6.8%) |       |
| 2  | of   | 36215 | (3.6%) |       |
| 3  | and  | 28793 | (2.8%) |       |
| 4  | to   | 26069 | (2.6%) |       |
| 5  | a    | 23383 | (2.3%) |       |
| 6  | in   | 21309 | (2.1%) |       |
| 7  | that | 10531 | (1.0%) |       |
| 8  | is   | 10004 | (1.0%) |       |
| 9  | was  | 9759  | (1.0%) |       |
| 10 | he   | 9494  | (0.9%) | 24.2% |
| 11 | for  | 9449  | (0.9%) |       |
| 12 | it   | 8703  | (0.9%) |       |
| 13 | with | 7249  | (0.7%) |       |
| 14 | as   | 7215  | (0.7%) |       |
| 15 | his  | 6960  | (0.7%) |       |
| 16 | on   | 6721  | (0.7%) |       |
| 17 | be   | 6340  | (0.6%) |       |
| 18 | at   | 5356  | (0.5%) |       |
| 19 | by   | 5317  | (0.5%) |       |
| 20 | i    | 5156  | (0.5%) | 30.8% |
| 21 | had  | 5120  | (0.5%) |       |
| 22 | this | 5119  | (0.5%) |       |
| 23 | not  | 4569  | (0.4%) |       |
| 24 | are  | 4360  | (0.4%) |       |
| 25 | but  | 4357  | (0.4%) | 32.0% |

# 단어 빈도수(word frequency)의 활용

- 초·중·고 교과과정에서 단계별 필수 어휘 선정
- 사전 편찬에서 표제어 서정, 정의 어휘, 의미 나열 순서
- 기초영어 교육에 필요한 최소한의 어휘 목록
- 1950년대:  
A General Service List of English Words
- 2000년대:  
A New General Service List (NGSL)  
A New Academic Word List (NAWL)
- NGSL (기초 어휘 2800 단어):  
일반 언어 사용의 86%
- NAWL (학술 어휘 960 단어 추가):  
일반 언어 사용의 92%



# 콘코던스(Concordance)

- 콘코던스: 코퍼스에서의 어 '색인'이라는 말은 '코퍼스' 단어나 어구가 사용된 용례 문맥과 함께 보여주는 목록
- 역사적으로 16세기 Bible Concordance, 19세기 Shakespeare Concordance
- 'naked + NOUN' 구조에서 있는 명사는 어떤 것이 있을 코퍼스에서 가장 많이 나오 'naked + eye'는 어떤 문맥 쓰일까?
- 단어의 문맥이나 구조를 시 명확하게 제시
- 단어나 어구의 전형적 사용
- 학습자가 '언어 탐정(language detectives)'이 되어 용례 파악
- 언어를 '발견'하는 Data-driven

cut but lower clarity (SI2 ) looks identical to the latex . " # The smallest particulates , invisible to the that are far away and objects that are invisible to the : ( a ) the existence of realities invisible to the a generation of highly mobile nymphs that are invisible to the anisotropies in the CMB are extremely faint , invisible to the come into view , i.e. , what was invisible to the of Venice in 1609 , even an entity invisible to the more than a few microns in diameter , invisible to the Mars will look as large as the Full Moon to the " " Invisible to you , but not necessarily to the it the Garnet Star . Its color is noticeable to the , Mercury and Venus appear as a single object to the some ancient native Floridian . That much was obvious to the different nature opens itself to the camera than opens to the , the Magellanic Clouds are the most impressive sights to the brightest and closest spiral in the sky . Visible to the an isosceles triangle of 5th-magnitude " stars " visible to the light-years , the eerie , glowing object is visible to the easy . Unlike human fingerprints , which are visible to the , NGC 3532 in Carina , is easily visible to the of Beta Scuti and on good nights is visible to the an emission nebula 800 light-years across that is visible to the help carry them farther . Thrips are barely visible to the of soymilk . Is the International Space Station visible to the a camera . None of these patterns is visible to the do not need a carrier beam to be visible to the would appear much brighter and would be easily visible to the thinking the laser -- whose light is not visible to the first report of a comet that was not visible to the the most dramatic celestial performance visible to the

naked eye but sells for just \$6,000 or so . The garb . naked eye , pose the greatest risk because they can be absorbed into naked eye with greater ease . # PHOTO (BLACK & WHITE ) naked eye ; ( b ) the ceaseless and typically cyclical dynamism of naked eye . These scurry about until they find a suitable spot in naked eye and even to most telescopes . In 1992 , the Cosmic naked eye is visible to the infra-red . Bernzy grasps the print with naked eye , such as an enemy ship on the horizon , would naked eye . Its core was three meters from where they stood . naked eye . " This statement first appeared in an e-mail during naked eye . Some unknown adversary , some secret enemy " " That naked eye and striking in binoculars . After Cepheus , the next stellar naked eye as they pass closer than at any time in the past naked eye . The maxilla , the still-intact upper jawbone , bore clear naked eye -- if only because an unconsciously penetrated space is naked eye , and are alone worth the trip to Australia . The naked eye , the galaxy is easy to locate . Begin at the naked eye . However , the apex of this triangle is not a naked eye on any clear winter night . Binoculars reveal bright stars naked eye and can be rapidly categorized by their swirl patterns , gun naked eye . By now the North would seem to be in full naked eye . M11 has been known as the Wild Duck Cluster since naked eye . A small telescope shows the loops of gas that look naked eye , and they move rapidly when bothered . If you blow naked eye in the night skies ? Jason McMurtrey Idaho Falls , Idaho naked eye or even easily described , but after studying thousands of naked eye as do the CO 2 systems . The CO 2 beam naked eye . " Try to glimpse this elusive galaxy on and exceptionally naked eye -- was turned off . # In fact , the laser naked eye at the time of its discovery . This catapulted her to naked eye , and leaving a subtle reminder of the fragility of life

# 콜로케이션(collocation)

- 콜로케이션: '연어관계'라고 하며, '텍스트 내의 일정한 공간에서 두 개 이상의 단어가 공

1. The mite is just visible to the **naked eye** and feeds on honey bees and their  
 2. they are not always visible to the **naked eye**. They are revealed in their million  
 3. when they are easily visible to the **naked eye** and can be fed in captivity. At al  
 4. the other stars that are visible to the **naked eye** lie within a few hundred light year  
 5. not a single one is visible to the **naked eye**. Another 15 per cent of stars are  
 6. It should be visible to the **naked eye** as a fuzzy patch ( without a tail )  
 7. the human egg is just visible to the **naked eye**. The three key cell structures are  
 8. though Zeta is always visible with the **naked eye** under even reasonable conditions  
 9. it is dimly visible with the **naked eye** in the guise of a slightly brighter  
 10. both are visible with the **naked eye**, not far from the Scorpion's stin  
 11. It is visible with the **naked eye**, and its position close to Eta and  
 12. It is clearly visible with the **naked eye**, close to Nu ( 4.5 ). Its distance  
 13. It is easily visible with the **naked eye**, and binoculars give it a vaguely  
 14. The faintest stars normally visible with the **naked eye** are of magnitude 6, and binoculars

- 'A word is keeps' -

- Adjacent  
 단어  
 근치  
 표  
 드  
 슨  
 한  
 관

- Window  
 의  
 한  
 분  
 석  
 기  
 준  
 으  
 로  
 공  
 간  
 에  
 서

-5 -4 -3

- 국내 매스  
 기  
 사  
 가  
 나  
 추  
 출  
 하  
 는  
 영  
 어  
 로  
 는

# 이런 정보가 사전 편찬에 어떻게 활용되는가?

## • 코퍼 Engli

## - 표제

## - 단어

## - 의미

## - 예문

## • 단어

### dry ◆◆◆◇

#### 1 dry ☞ drier or dryer ☞ driest ☞

If something is **dry**, there is no water or moisture on it or in it.

Clean the metal with a soft dry cloth.

Pat it dry with a soft towel.

Once the paint is dry, apply a coat of the red ochre emulsion paint.

The path was dry and slithery from the drought.

ADJ-GRADED

≠ wet, damp

#### • dryness ☞

...the parched dryness of the air.

N-UNCOUNT

#### 2 dry ☞ or dries ☞ drying ☞ dried ☞

When something **dries** or when you **dry** it, it becomes dry.

The washing might dry outside today, the sun's shining.

Leave your hair to dry naturally whenever possible.

Wash and dry the lettuce.

Liz laughed again, got up from the water and began to dry herself.

V-ERG

#### 3 dry ☞ or dries ☞ drying ☞ dried ☞

When you **dry** the dishes after a meal, you wipe the water off the plates cups, knives, pans, and other things when they have been washed, using a cloth.

Mrs. Madrigal picked up a towel and began drying dishes next to her daughter.

VB

= wipe

#### + dry up ☞ or dries up; drying up ☞ dried up

**Dry up** means the same as **dry**. (BRIT)

He got up and stood beside Julie, drying up the dishes while she washed.

PHR-V

#### 4 dry ☞ drier or dryer ☞ driest ☞

If you say that your skin or hair is **dry**, you mean that it is less oily than or not as soft as normal.

Nothing looks worse than dry, cracked lips.

Dry hair can be damaged by washing it too frequently.

My skin's been getting a little dry recently.

ADJ-GRADED

≠ greasy

#### • dryness ☞

Dryness of the skin can also be caused by living in centrally heated homes and offices.

N-UNCOUNT

#### 5 If the weather or a period of time is **dry**, there is no rain or there is much less rain than average.

Exceptionally dry weather over the past year had cut agricultural production...

ADJ

≠ wet

#### 6 A **dry** place or climate is one that gets very little rainfall.

...a hot, dry climate where the sun is shining all the time.

ADJ usu ADJ n

= arid ≠ wet

#### • dryness ☞

He was advised to spend time in the warmth and dryness of Italy.

N-UNCOUNT

#### 7 In the **dry** means in a place or at a time that is not damp, wet, or rainy. (mainly BRIT)

Such cars, however, do grip the road well, even in the dry.

N-SING the N, usu in N

≠ wet

#### 8 If a river, lake, or well is **dry**, it is empty of water, usually because of hot weather and lack of rain.

ADJ

#### 9 If an oil well is **dry**, it is no longer producing any oil.

ADJ usu v-link ADJ

#### 10 If your mouth or throat is **dry**, it has little or no saliva in it, and so feels very unpleasant, perhaps because you are tense or ill.

His mouth was dry, he needed a drink...

ADJ usu v-link ADJ

#### • dryness ☞

Symptoms included frequent dryness in the mouth.

N-UNCOUNT usu with supp

#### 11 If someone has **dry** eyes, there are no tears in their eyes; often used with negatives or in contexts where you are expressing surprise that they are not crying.

There were few dry eyes in the house when I finished...

ADJ

≠ moist

#### 12 If a country, state, or city is **dry**, it has laws or rules which forbid anyone to drink, sell, or buy alcoholic drink. (INFORMAL)

Gujarat has been a totally dry state for the past thirty years.

ADJ

= teetotal

#### 16 dry ☞ drier or dryer ☞ driest ☞

**Dry** humour is very amusing, but in a subtle and clever way.

Though the pressure Fulton is under must be considerable, he has retained his dry humour.

Mr Brooke is renowned for his dry wit.

ADJ-GRADED: usu ADJ n approval

= witty

#### • drily ☞

As Rossini drily observed, 'Wagner has lovely moments but awful quarters of an hour.'

...a frank and drily witty woman.

ADV: ADV with v, ADV adj

#### • dryness ☞

It has a wry dryness you won't recognise.

N-UNCOUNT

#### 17 dry ☞ drier or dryer ☞ driest ☞

If you describe a voice as **dry**, you mean that it is cold or dull, and does not express any emotions. (WRITTEN)

When he crept back to his desk, he heard the dry voice of Father Laurence.

ADJ-GRADED

= thin ≠ rich

#### • drily ☞

'Possible,' I said drily, 'but not likely.'

ADV: ADV with v

#### 18 dry ☞ drier or dryer ☞ driest ☞

If you describe something such as a book, play, or activity as **dry**, you mean that it is dull and uninteresting.

My eyelids were drooping over the dry, academic phrases.

A lot of the work was very dry and boring in Westminster.

ADJ-GRADED disapproval

= dull, boring

#### 19 dry ☞

**Dry** bread or toast is plain and not covered with butter or jam.

For breakfast, they had dry bread and tea.

ADJ: ADJ n

= unbuttered

#### 20 dry ☞

**Dry** sherry or wine does not have a sweet taste.

...a glass of chilled, dry white wine.

ADJ

≠ sweet

# 코퍼스 분류

---

- Online Corpora

- English Corpora
- British National Corpus
- The Michigan Corpora
- ...

- Offline corpora (self-compiled)

- 영어학습자 코퍼스
- ESP 코퍼스
- 문학작품 코퍼스 (via [Project Gutenberg](#))
- ...

# 온라인 코퍼스에 AI 분석 기능 도입



## English-Corpora.org

corpora PDF guides videos related resources users my account upgrade help

Overview: [brief](#) | [detailed](#) 

Now available: AI/LLM insights integrated into corpus results

These are the most widely used online corpora, and they serve many different purposes for teachers and researchers at universities throughout the world. In addition, the corpus data (e.g. full-text, word frequency) has been employed by a wide range of companies in many different fields, especially technology and language learning.

The links below are for the free online interface. You can also purchase and download  the corpora for use on your own computer.

| Corpus                                         | Download                                                                            | # words       | Dialect      | Time period    | Genre(s)         |
|------------------------------------------------|-------------------------------------------------------------------------------------|---------------|--------------|----------------|------------------|
| News on the Web (NOW)                          |    | 24.3 billion+ | 20 countries | 2010-yesterday | Web: News        |
| iWeb: The Intelligent Web-based Corpus         |    | 14 billion    | 6 countries  | 2017           | Web              |
| Global Web-Based English (GloWbE)              |    | 1.9 billion   | 20 countries | 2012-2013      | Web (incl blogs) |
| Wikipedia Corpus                               |    | 1.9 billion   | (Various)    | 2014           | Wikipedia        |
| Coronavirus Corpus                             |   | 1.5 billion   | 20 countries | 2020-2023      | Web: News        |
| Corpus of Contemporary American English (COCA) |  | 1.0 billion   | American     | 1990-2019      | Balanced         |
| Corpus of Historical American English (COHA)   |  | 475 million   | American     | 1820-2019      | Balanced         |
| The TV Corpus                                  |  | 325 million   | 6 countries  | 1950-2018      | TV shows         |
| The Movie Corpus                               |  | 200 million   | 6 countries  | 1930-2018      | Movies           |
| Corpus of American Soap Operas                 |  | 100 million   | American     | 2001-2012      | TV shows         |

- Classifying and categorizing collocates
- Classifying and categorizing collocates (with PoS: COCA, iWeb)
- Classifying and categorizing phrases
- Comparing two words (via collocates)
- Comparing two genres, time periods, and dialects (lists)
- Comparing genres, time periods, and dialects (all sections)
- Comparing genres, time periods, and dialects (charts)
- Analyzing KWIC/concordance lines
- Generating words and phrases for topics, translations, and rephrasing

# Integrative corpus-AI analysis ([English Corpora](#))

---

- **Classifying and categorizing collocates (COCA)**
  - sheer + NOUN, ADJ/NOUN + fatigue
- **Comparing two words (COCA)**
  - stick to/adhere to, cure/heal
- **Analyzing KWIC/concordance lines (NOW)**
  - naked eye, NOUN set in
- **Comparing time periods (COHA)**
  - spilled/spilt, maybe/perhaps, centre/center
- **Comparing time periods & categorizing (COHA)**
  - [regret] ~ing

Classifying and categorizing collocates

Classifying and categorizing collocates (with PoS: COCA, iWeb)

Classifying and categorizing phrases

Comparing two words (via collocates)

Comparing *two* genres, time periods, and dialects (lists)

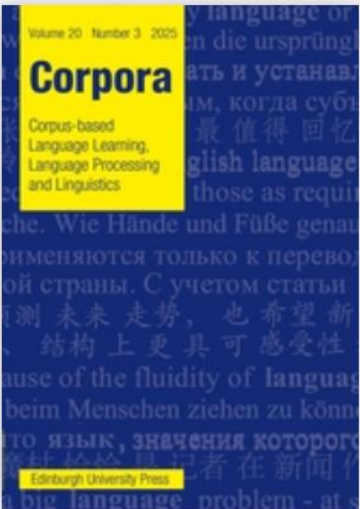
Comparing genres, time periods, and dialects (all sections)

Comparing genres, time periods, and dialects (charts)

Analyzing KWIC/concordance lines

Generating words and phrases for topics, translations, and rephrasing

# 학습자코퍼스와 제2언어 연구: LCR in Korea



**Table of Contents**  
Volume 17 | Issue  
Supplement | October 2022  
Learner Corpora Research in the Asia  
Pacific Region  
ISSN: 1749-5032  
eISSN: 1755-1676  
3 issues per year: April, August & November

 **Free access** | October 2022  
English learner corpora and research in Korea  
Heokseung Kwon  
<https://doi.org/10.3366/cor.2022.0244> • pp 5-22

[Preview abstract](#) ▾      [Citation](#) | [Full text](#) | [References](#) |

|           |                                                                                                                                                                                                                                                                 |
|-----------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 2001–2005 | Kim (2002)<br>Lee (2004)                                                                                                                                                                                                                                        |
| 2006–2010 | Lee (2006)<br>Oh (2007)<br>Song (2008)<br>Oh (2009)<br>Yu and Yoo (2010)                                                                                                                                                                                        |
| 2011–2015 | Yoon and Yoo (2011)<br>Jung and Yoo (2011)<br>Cho and Yoo (2012)<br>Choi and Yoo (2012)<br>Park (2013a)<br>Park (2013b)<br>Cho and Shin (2014)<br>Oh (2014)<br>Kim (2014)<br>Chang (2015)<br>Jin (2015)<br>Lah and Yoo (2015)<br>Lee, Yook, Lee and Park (2015) |
| 2016–2020 | Kim (2016)<br>Lee, Yook and Park (2016)<br>Choung and Oh (2017)<br>Chang (2018)<br>Shin (2018)<br>Hong (2019)<br>Ishikawa (2019)<br>Sung and Kim (2020)<br>Kim (2020)                                                                                           |
| 2021      | Jo and Oh (2021)                                                                                                                                                                                                                                                |

|           |                                                                                                                                                                              |
|-----------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 2001~2005 | Lee (2005)                                                                                                                                                                   |
| 2006~2010 | Yoon (2006)<br>Lee (2006)<br>Kweon (2008)<br>Song (2008)<br>Huh (2008)<br>Chon and Shin (2009)<br>Oh (2009a)<br>Oh (2009b)                                                   |
| 2011~2015 | Kim (2011)<br>Oh and Kang (2011)<br>Shin (2011)<br>Park (2012)<br>Kim (2012)<br>Chang (2013)<br>Sung and Yoo (2013)<br>Oh (2014)<br>Cho (2014)<br>Park (2015)<br>Kwon (2015) |
| 2016~2020 | Park (2017)<br>Sung (2017)<br>Park (2017)<br>Park (2018)<br>Kim (2018)<br>Pyo (2018)<br>Lee and Jun (2018)<br>Lee (2019)<br>Kim (2019)<br>Ryu (2020)<br>Yu and Oh (2020)     |
| 2021      | Pyo (2021)                                                                                                                                                                   |

# 학습자코퍼스와 제2언어 연구: LCR in East Asia

---

Kwon (2027): Corpus-AI integrative approach to Korean EFL learner corpus research



---

## Proposal for AsiaTEFL Book Series

**Title:** Using Corpora and AI for Second Language Research in East Asia

**Editors:** Isaiah WonHo Yoo, Dilin Liu, Yaochen Deng, and Yukio Tono

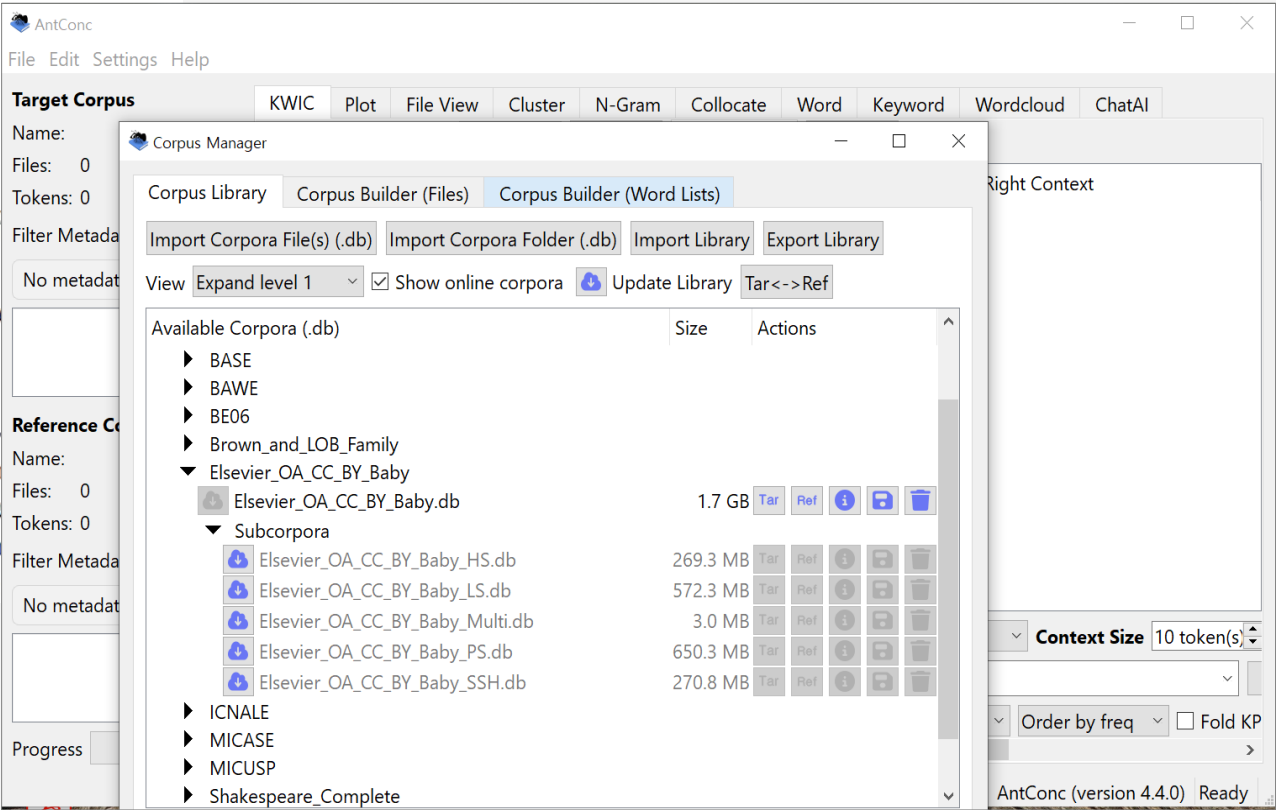
# ESP 연구

ESP Review Vol. 8, No. 2 Summer 2026  
<https://doi.org/10.23191/espkor.2026.8.2.73>

## ESP 연구: 코퍼스에서 AI와의 융합으로

Kwon, Heokseung. (2026). ESP research: From corpora to AI integration. *ESP Review* 8(2), 73-100.

This article examines the historical development of ESP research from corpus approaches to recent AI-integrated methodologies and explores future directions for studies in the age of artificial intelligence. It first reviews how corpus linguistics has been closely connected with ESP research since the 1990s through advances in computing technology, large-scale corpora, and corpus analysis tools. The study then introduces



Q & A

---

감사합니다!

