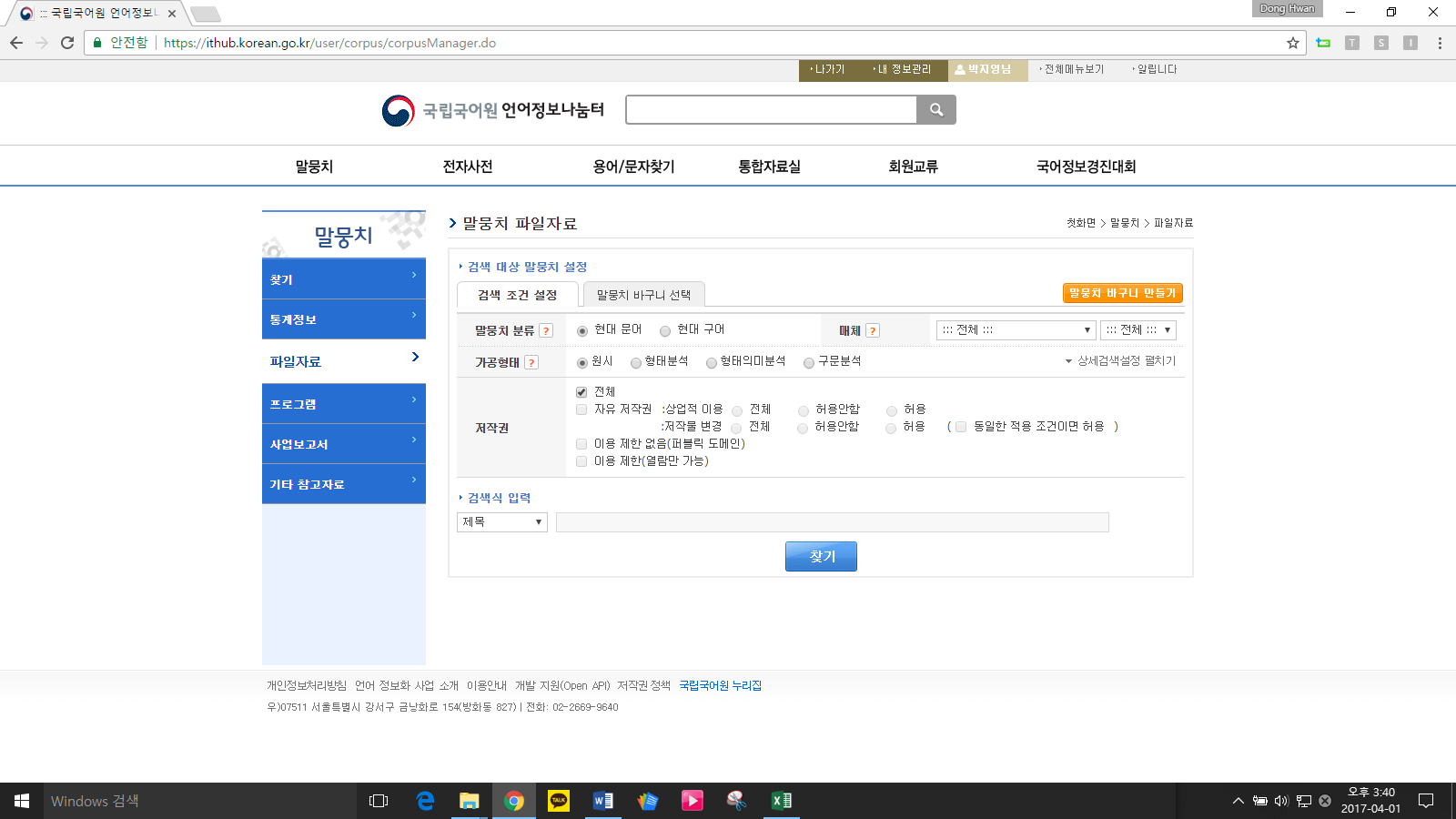
한국에서는 1988년부터 구축되어 온 연세대학교의 연세 한국어 말뭉치가 대표적이다. 한국과학기술원, 고려대학교, 국립국어연구원 등에서도 본격적으로 말뭉치를 구축해 오고 있으며, 1998년부터 ‘21세기 세종 계획: 국어정보화 추진 중장기 사업’의 일환으로 우리나라에서도 본격적으로 국가 말뭉치가 구축되기 시작하였다.

1. 연세 한국어 말뭉치(1987-): 1960년대 이후의 한국어, 4,200만 어절, 연세 한국어 사전을 만드는 데 활용. (연세대학교 언어정보연구원)
2. 고려대학교 한국어 말모둠(1995): 1970-90년대 한국어: 1,000만 어절 (고려대학교 민족문화연구소)
3. 국립국어원 말뭉치(1992-): 1400년대 이후의 한국어, 6,800만 어절, 표준국어대사전을 만드는 데 활용 (국립국어원)

말뭉치에는 그 특성에 따라 여러 종류가 있다.

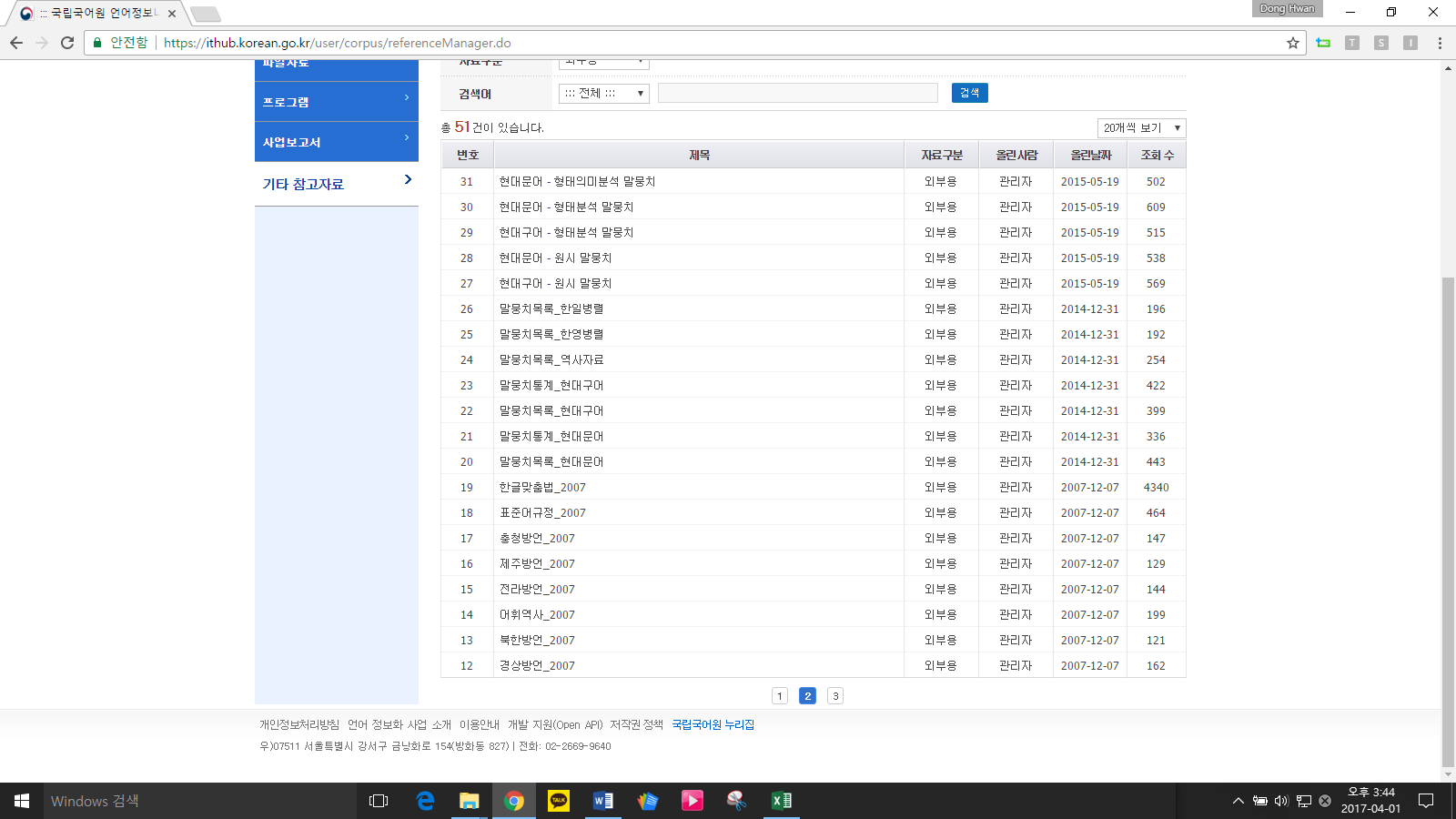
1. 용도에 따라: 균형 말뭉치(balanced/core corpus), 학습 말뭉치(training corpus), 학습자 말뭉치(learner’s corpus), 방언 말뭉치(dialect corpus)
2. 가공 정도에 따라: 원시 말뭉치(raw corpus), 주석 말뭉치(tagged/annotated corpus), 분석 말뭉치(analyzed corpus)
3. 시대에 따라: 공시적 말뭉치(synchronic corpus), 역사 말뭉치(historical/diachronic corpus)
4. 변화여부에 따라: 정적 말뭉치(static corpus), 동적 말뭉치(dynamic/monitor corpus)



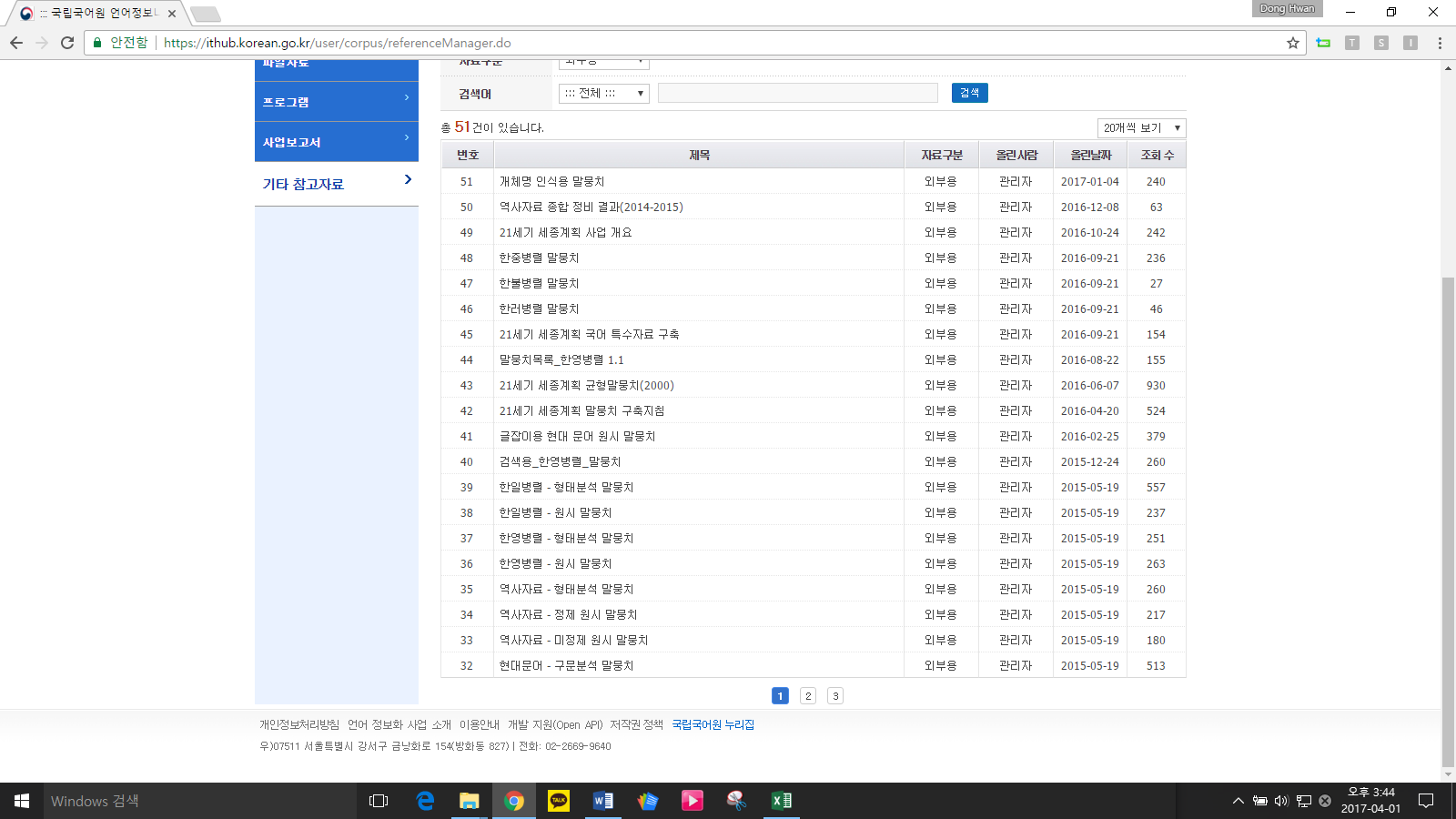
출처: 국립국어원

말뭉치 분류: 현대 문어, 현대 구어

가공형태: 원시, 형태분석, 형태의미분석, 구문분석

출처: 국립국어원 홈페이지(https://ithub.korean.go.kr/user/corpus/referenceManager.do)

1. 경상방언
2. 북한방언
3. 어휘역사
4. 전라방언
5. 제주방언
6. 충청방언
7. 표준어규정
8. 한글맞춤법
9. 말뭉치목록\_현대문어
10. 말뭉치통계\_현대문어
11. 말뭉치목록\_현대구어
12. 말뭉치목록\_현대구어
13. 말뭉치목록\_역사자료
14. 말뭉치목록\_한영병렬
15. 말뭉치목록\_한일병렬
16. 현대구어 – 원시 말뭉치
17. 현대문어 – 원시 말뭉치
18. 현대구어 – 형태분석 말뭉치
19. 현대문어 – 형태분석 말뭉치
20. 현대문어 – 형태의미분석 말뭉치



출처: 국립국어원 홈페이지(<https://ithub.korean.go.kr/user/corpus/referenceManager.do>)

1. 현대문어 – 구문분석 말뭉치
2. 역사자료 – 미정제 원시 말뭉치
3. 역사자료 – 정제 원시 말뭉치
4. 역사자료 - 형태분석 말뭉치
5. 한영병렬 – 원시 말뭉치
6. 한영병멸 – 형태분석 말뭉치
7. 한일병렬 – 원시 말뭉치
8. 한일병렬 – 형태분석 말뭉치
9. 검색용\_한영병렬\_말뭉치
10. 글잡이용 현대 문어 원시 말뭉치
11. 21세기 세종계획 말뭉치 구축지침
12. 21세기 세종계획 균형말뭉치(2000)
13. 말뭉치목록\_한영병렬 1.1
14. 21세기 세종계획 국어 특수자료 구축
15. 한러병렬 말뭉치
16. 한불병렬 말뭉치
17. 한중병렬 말뭉치
18. 21세기 세종계획 사업 개요
19. 역사자료 종합 정비 결과(2014-2015)
20. 개체명 인식용 말뭉치