

말뭉치 문체론의 관점에서 본 한국어의 발생 문법 양상¹⁾

- ‘-거든’의 공시적·통시적 실증적 탐구를 사례로 -

제조현(연세대학교)

이 글은 크게 두 가지 목적을 목표로 한다. 첫째, 발생 문법(잠정 문법, Emergent Grammar, 浮現文法)의 정의와 유형을 귀납적으로 정리하고, 그 유형 중 하나인 국어학계에서 주목받아 온 ‘종결어미화된 연결어미’(혹은 탈종속화) 현상의 연구사와 주요 쟁점을 밝히는 것이다. 둘째, ‘문체 변이(Stylistic Variations)’가 사회언어학과 변이언어학에서 하위 변인으로 작용하며, ‘생략, 도치, 문법화 등 다양한 요인으로 형성된 연결어미의 의미 확장 과정’에 실제로 영향을 미쳤는지를 탐구하는 데 있다. 이를 위해 ‘-거든’을 대상으로 공시적·통시적 말뭉치에서 추출한 용례를 분석하며, 문체 변인과 연결어미의 의미확장 간의 통계적 상관관계 및 교호작용을 다변량분석과 혼합효과 로지스틱 회귀분석을 통해 실증적으로 규명하고자 한다.

이 글의 구성은 다음과 같다. 2장에서는 예비적 검토를 통해 다음 두 가지를 다룬다. 첫째, 발생 문법과 사용-기반 이론(usage-based theory)의 문법적 관점을 살펴보고, 한국어에서 ‘(제3) 매체인 메신저 대화’와 ‘구어’의 용례를 분석함으로써 잠정 문법 현상의 유형을 규명한다. 아울러, ‘매체(media)’, ‘문체(style)’, ‘사용역(register)’, ‘장르(genre)’ 등의 학술적 용어와 변량(variables)이라는 개념을 연구에 도입한다. 둘째, ‘구어 문법’의 관점에서 연결어미와 종결어미의 체계를 재검토하며, ‘종결어미화된 연결어미’의 연구사를 정리하고 향후 과제를 논의한다.

3장에서는 본 연구에서 활용한 현대국어 균형 말뭉치와 시기별·장르별·유형별로 구성된 역사 자료 말뭉치의 구성을 먼저 소개할 것이다. 이후, 말뭉치에서 추출된 용례의 기술 통계(descriptive statistics) 결과를 보고할 것이다. 3.2절에서는 연구

1) 발표 자료와 PPT는 아래 주소에서 다운로드할 수 있습니다:
https://github.com/qizhaoxian/20241221_PPT

방법론을 다룰 예정으로, R을 활용한 말뭉치언어학의 다변량분석 방법과 전통적인 말뭉치언어학의 연구 방법을 비교함으로써, 본 연구의 목적을 달성하기 위해 채택한 연구 방법을 상세히 설명할 것이다.

4장에서는 현대국어에서 ‘-거든’을 대상으로 수행한 실험 결과를 다룬다. 대중대응분석(MCA)과 혼합효과 로지스틱 회귀분석 등을 활용하여, 텍스트 장르에 따른 접속 위치와 문말 위치에서의 ‘-거든’ 사용 양상을 고찰할 것이다.

5장에서는 ‘-거든’의 통시적 실험 결과를 다차원적 척도분석(MDS analysis)을 통해 시각화하고, 이를 기반으로 ‘-거든’의 통시적 의미변화를 분석한다. 아울러 문체가 의미변화에 미치는 영향을 논의할 것이다.

부록 1. 각 변수(Variable)와 변수 수준(Variable Level)의 주석 기준, 그리고 예시 화면²⁾

<1> 사용역(register): Regi	
구어(spoken language): Oral	0
문어(written language): Writ	1
<2> 장르(genre): Genre	
공적 대화 (public conversation): PuCo	0
공적 독백 (public monologue): PuMo	1
사적 대화 (private conversation): PrCo	2
사적 독백 (private monologue): PrMo	3
준구어 (semi-spoken language): SeSp	4
소설 (novel): Novl	5
잡지 (magazine): Magz	6
학술 교양 (academic and educational materials): AcEd	7
국회 회의록 (parliamentary minutes): Parl	8
신문 (newspaper): News	9
<3> 담화 기능(discourse function): DFnc	
DiscFA	0
DiscFB	1
DiscFC	2
DiscFD	3
DiscFE	4
DiscFF	5
<4> 결합어(construction collocation): CoCo	

2) 영어: 대중대응분석 등 방법의 주석 기준; 숫자: 혼합효과 로지스틱 회귀분석의 주석 기준

<5> 의미운율(semantic prosody): SePr

적극적(positive): Posi	0
중립적(neutral): Neut	1
소극적(negative): Nega	2

<6> 인접 범주(colligation): Collig

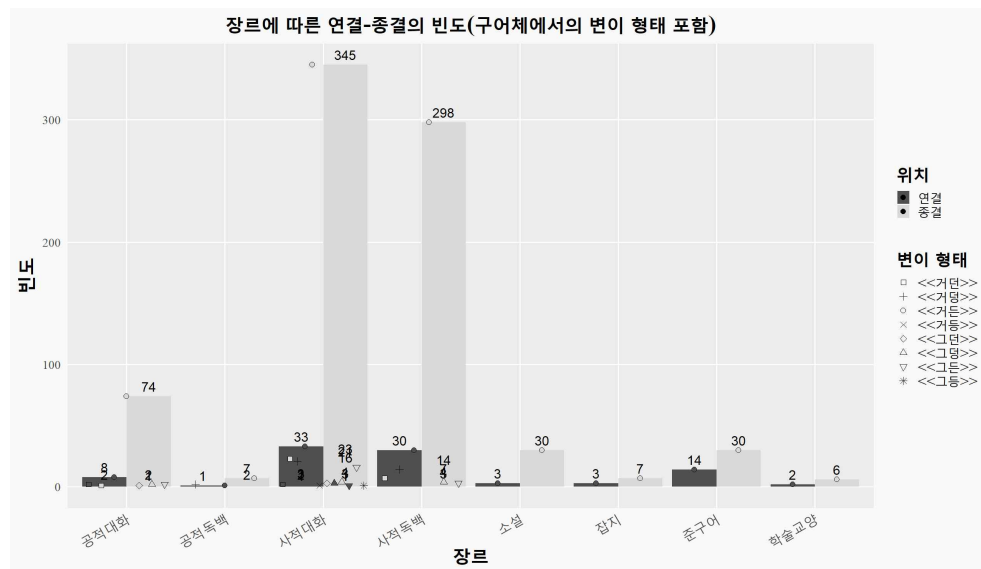
동사(verb): Verb	0
보조용언(auxiliary predicate): AuxP	1
지정사(copula): Copl	2
형용사(adjective): Adjv	3

<7> 구어 변이형(spoken variant form): SpVa

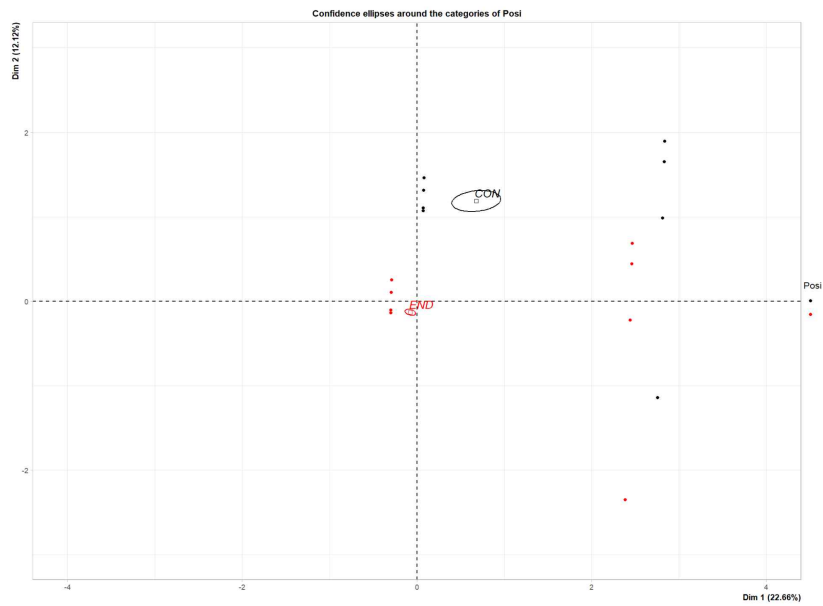
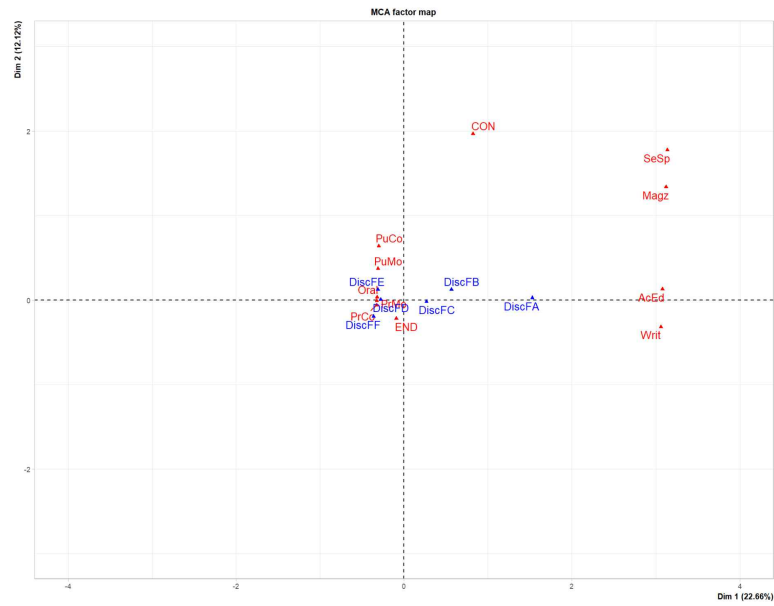
<8> 위치(position): Posi

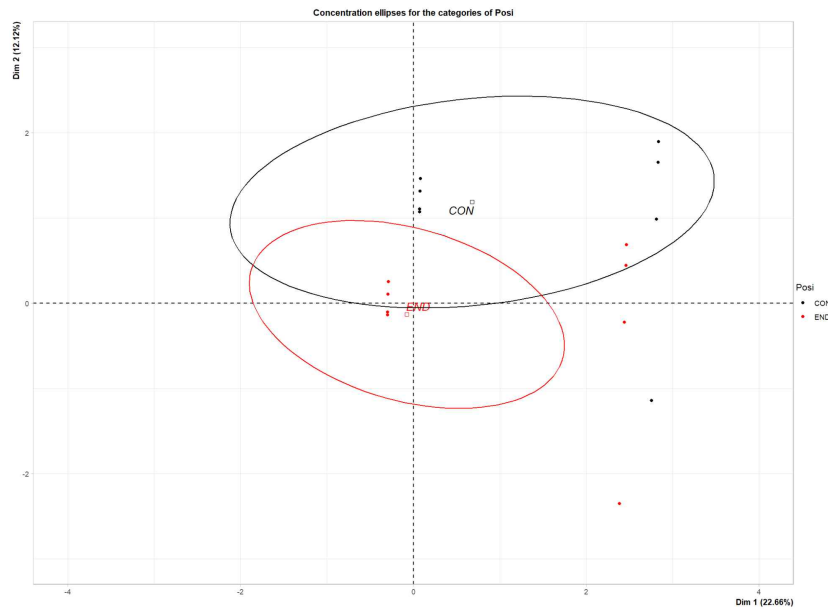
연결(connection): CON	0
종결(ending): END	1

부록 2. 기술 통계의 결과



부록 3. 다중대응분석(MCA)의 시각화 결과





부록 4. 혼합효과 로지스틱 회귀분석의 결과

