

규칙, 제약, 최대 엔트로피

Rule, constraint and maximum entropy

전종호, 서울대학교
Jongho Jun, Seoul National University

생성음운론 (Generative Phonology)

- Chomsky, Noam & Morris Halle (1968) *The Sound Pattern of English* (SPE)
- 발전 시기별 주류 이론
 1. 규칙 기반 (Rule-based phonology)
 2. 최적성 이론 (Optimality Theory: OT)

생성음운론 (Generative Phonology)

- Chomsky, Noam & Morris Halle (1968) *The Sound Pattern of English* (SPE)
- 발전 시기별 주류 이론
 1. 규칙 기반 (Rule-based phonology)
 2. 최적성 이론 (Optimality Theory: OT)



비범주적, 양적 양상 (quantitative patterns)

통계적 문법 모형

범주적 양상

생성음운론 (Generative Phonology)

- Chomsky, Noam & Morris Halle (1968) *The Sound Pattern of English* (SPE)
- 발전 시기별 주류 이론
 1. 규칙 기반 (Rule-based phonology)
 2. 최적성 이론 (Optimality Theory: OT)
 3. 최대 엔트로피 조화 문법 (Maximum Entropy Harmonic Grammar: MaxEnt)

규칙 기반 (Rule-based phonology)

- 경험적 자료 기술
- 이론적 문법 분석

규칙 기반 (Rule-based phonology)

예) 영어 어말 t/d 삭제: 'west (side)' wesØ

- 자료 기술 단어 말 치조 파열음은 다른 자음 뒤에서 삭제된다.
- 문법 분석 문법 = 규칙 (+ 규칙 순서) (Chomsky & Halle 1968)

• Final t/d deletion: $\left[\begin{array}{l} \text{-sonorant} \\ \text{-continuant} \\ \text{+coronal} \end{array} \right] \rightarrow \emptyset / \text{C} _$

규칙 기반 (Rule-based phonology)

- 샘플 도출 (Sample derivation)

기저형 (Underlying Representation)

west

Final t/d deletion

∅

표면형 (Surface Representation)

wes

최적성 이론 (Optimality Theory: OT)

- 자료 기술
- 문법 분석 문법 = 제약 + 랭킹 (Prince & Smolensky 1993/2004)

최적성 이론 (Optimality Theory: OT)

- 문법 분석 문법 = 제약 + 랭킹 (Prince & Smolensky 1993/2004)

제약

- *CT '치조 파열음으로 끝나는 단어 말 자음군 회피'
- MAX '삭제 금지'

제약 랭킹: *CT >> MAX

최적성 이론 (Optimality Theory: OT)

- 분석표 (Tableau)

west	*CT	MAX
a.  wes Ø		*
b. west	*	

최대 엔트로피 조화문법 (MaxEnt)

전통적인 음운론 연구 대상

- 범주적 (categorical) 현상: O
- 비범주적 (gradient) 현상: X
 - 변이 (variation), 예외 (exception)

→ 비범주적 현상 \approx 범주적 현상 (Coetzee & Pater 2011)

최대 엔트로피 조화문법 (MaxEnt)

- 영어 단어 말 t/d 삭제: **수의적 (optional)**
 - wes[t] ~ wes[∅] side

최대 엔트로피 조화문법 (MaxEnt)

- 영어 어말 t/d 삭제: 수의적 (optional)

• wes[t] ~ wes[∅] side

- 환경에 따라 삭제 비율 다름

모음 앞

휴지 (pause) 앞

<<

자음 앞

e.g. west end

west ##

west side

cf. 범주적 자음 삭제 현상: 주로 자음 앞에서만 발생

- 문법 분석: 발생 비율 포착

최대 엔트로피 조화문법 (MaxEnt)

- 환경에 따른 차이: pre-V, pre-pause << Pre-C
- 방언에 따른 차이 (Coetzee & Pater 2008 ms: 9 (10))

t/d 삭제 비율 (%)	모음 앞	휴지 앞	자음 앞
Washington, DC	29	73	76
Chicano English	45	37	62
Jamaican English	63	71	85
Tejano English	25	46	62
Trinidadian English	21	31	81
Philadelphia English	38	12	100

최대 엔트로피 조화문법 (MaxEnt)

- 문법 = 제약 + 가중치 (weight) (Goldwater & Johnson 2003, Hayes & Wilson 2008)
- 최대 엔트로피 조화문법에서는 제약의 강도를 숫자 가중치(weight)로 표시하며, 표면 형태가 위배하는 제약의 가중치 및 위배 횟수에 따라 그 형태의 발생 확률이 부여된다.

최대 엔트로피 조화문법 (MaxEnt)

- 수의적 영어 어말 t/d 삭제

e.g. wes~~∅~~ ~ west

가중치

발생 확률

	1	1	H	e^H	P
west	*CT	MAX			
a. wes ∅		-1	-1	0.36	0.5
b. west	-1		-1	0.36	0.5

최대 엔트로피 조화문법 (MaxEnt)

- 수의적 영어 어말 t/d 삭제

e.g. wes~~∅~~ ~ west^t

가중치

발생 확률

	2	1	H	e ^H	P
west	*CT	MAX			
a. wes ∅		-1	-1	0.36	0.73
b. west ^t	-1		-2	0.14	0.27

최대 엔트로피 조화문법 (MaxEnt)

▪ 수의적 영어 어말 t/d 삭제 (Chicano English)

t/d 삭제 비율 (%)	모음 앞	휴지 앞	자음 앞
Chicano English	45	37	62

▪ 제약 (Coetzee & Pater 2011)

*CT	t/d로 끝나는 자음군 회피
MAX	삭제 금지
MAX-Pre-V	모음 앞 삭제 금지
MAX-Final	휴지(pause) 앞 삭제 금지

최대 엔트로피 조화문법 (MaxEnt)

- 수의적 영어 어말 t/d 삭제 (Chicano English)

t/d 삭제 비율 (%)	모음 앞	휴지 앞	자음 앞
Chicano English	45	37	62

	?	?	?	?	
west (side)	*CT#	MAX-PRE-V	MAX-Fin	MAX	P
a. wes o				-1	?
b. west	-1				?

최대 엔트로피 조화문법 (MaxEnt)

- MaxEnt 문법 = 제약 + 제약 가중치
- 제약 가중치 찾기 \approx 문법 학습
- 정확한 제약 가중치 계산을 위해서는 프로그램 사용
- 문법 학습 프로그램은 인간 화자의 발화 자료를 학습 자료로 삼아서 문법 습득을 수행함으로써 인간의 문법 습득을 모델링

최대 엔트로피 조화문법 (MaxEnt)

- 정확한 제약 가중치 계산을 위해 프로그램 사용: e.g. Maxent Grammar Tool
- 학습 자료 (training data) → 문법 학습 시뮬레이션

t/d 삭제 비율 (%)	모음 앞	휴지 앞	자음 앞
Chicano English	45	37	62

	?	?	?	?	
west (side)	*CT#	MAX-PRE-V	MAX-Fin	MAX	P
a. wes o				-1	?
b. west	-1				?

최대 엔트로피 조화문법 (MaxEnt)

- 정확한 제약 가중치 계산을 위해 프로그램 사용: e.g. Maxent Grammar Tool
- 학습 자료 (training data) → 문법 학습 → 테스트 시뮬레이션

t/d 삭제 비율 (%)	모음 앞	휴지 앞	자음 앞
Chicano English	45	37	62

	5.24	0.69	1.02	4.75	
west (side)	*CT#	MAX-PRE-V	MAX-Fin	MAX	P
a. wes Ø				-1	0.62
b. west	-1				0.38

- 규칙 기반
- 최적성 이론
- 최대 엔트로피 문법
- 요약

요약: 음운론 연구

1. 경험적 자료 기술
2. 이론적 문법 분석

1. 경험적 자료 기술

저자 직관



코퍼스/실험 → 통계

- 비범주적 자료/현상으로 연구 대상 확대

2. 이론적 문법 분석

1. 경험적 자료 기술

저자 직관



코퍼스/실험 → 통계

- 빈도주의 혼합 효과 모델(Mixed effects models)
→ 베이지언 (Bayesian) 모델

2. 이론적 문법 분석

1. 경험적 자료 기술

저자 직관



코퍼스/실험 → 통계

2. 이론적 문법 분석

규칙



OT 제약 & 랭킹



제약 & 가중치

샘플 도출 (Sample derivation)

→ 분석표 (Tableau)

→ 문법 학습 & 테스트 시뮬레이션

- 규칙 기반
- 최적성 이론
- 최대 엔트로피 문법
- 요약

감사합니다.

- Chomsky, Noam, & Morris Halle (1968) *The Sound Pattern of English*. New York: Harper & ROW.
- Coetzee, Andries & Joe Pater (2011) The place of variation in phonological theory. In J. Goldsmith, J. Riggle & A. Yu (eds.) *The Handbook of Phonological Theory*, 2nd edition, 401-434. Malden, MA and Oxford, UK: Blackwell.
- Goldwater, Sharon & Mark Johnson (2003). Learning OT constraint rankings using a Maximum Entropy model. In Jennifer Spenador, Andres Eriksson & Östen Dahl (eds.) *Proceedings of the Workshop on Variation within Optimality Theory*, 111-120. Stockholm: Stockholm University.
- Hayes, Bruce & Colin Wilson (2008) A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39, 379–440.
- Prince, Alan & Paul Smolensky (1993/2004) *Optimality theory: Constraint interaction in generative grammar*. Oxford, UK: Blackwell.