

“규칙, 제약, 최대 엔트로피(Rule, constraint and maximum entropy)”에 대한 토론문

이주희

경희대학교

juhee@khu.ac.kr

발표문 “규칙, 제약, 최대 엔트로피”에서는 지난 반세기 동안 이루어진 음운 이론의 성과와 현재를 변이형의 사례를 통하여 비교 분석한다. 생성문법에서는 분석 대상에 기저형과 표면형의 개념을 정의함으로써 적절한 구조를 부여하는 것을 최우선에 놓는데, 특히 음운론의 분야에서 기저형의 설정에 대한 문제는 뜨거운 논쟁의 대상이었다. Chomsky & Halle(1968)의 SPE 이후 지난 50년 동안 생성음운론이 눈부신 발전을 하였다는 점은 누구도 부정할 수 없는 사실이며, 이러한 규칙기반 음운론(Rule-based phonology)은 국어음운론의 연구와 방법에도 지대한 영향력을 미치게 되었다. 따라서 국어음운론에서 기저형에 대한 추상성에 대한 문제와 규칙의 자연성은 음운분석에서 중요한 개념이었으나, 규칙의 설정에 있어서 중간단계에서 표면형에 나타날 수 없는 추상적 형태를 임의로 도출하여 불투명성(opacity)을 야기하는 등의 문제점이 있으며, 실제 언어의 분석에서 기저형과 표면형 사이의 밀접성, 즉 충실성(faithfulness)을 제대로 지키지 못하는 경우가 빈번한 문제점이 있다.

최적성이론(Prince & Smolensky 1993)은 이러한 설명력에 대한 부담을 줄이고 입력부에 중심을 둔 이론으로서 하나의 생성방법을 이용하여 나타낼 수 있는 모든 형태를 만들어 내고 적절한 제약을 적용하여 이에 대한 위배가 최소로 적은 것을 최적(optimal) 형태로 인정하는 제약기반 음운론(Constraint-based phonology)이다. 90년대 초반에 대두된 최적성이론과 같은 제약기반 문법이 음운론의 연구와 방법에 다시금 큰 변화를 야기했지만 시간이 지나면서 연구자들은 여러 유형의 변이현상 중에서 주로 언어 내적인 요인, 다시 말해서 음운·형태적 이유로 통제되는 변이현상을 어떻게 다루어야 할지 활발한 논의를 한다. 가령 순위제약을 가정한 최적성이론에서 Hammond(1994)는 변이현상이 동점위배(tied violation)라는 개념을 통해 설명될 수 있다고 주장하지만, 최적성이론에서 제약이 범어적으로 결정되어 있는 것을 고려할 때 발생하는 문제점을 지적하지 않을 수 없다. 따라서 변이형을 보이는 후보형들은 주어진 제약에 대해 동점위배로 보일 수 있겠으나, 그 밖에 많은 하위 제약이 있을 수 있으며 변이형들이 동등하게 위배하지 않는 제약도 가능할 수 있다.

발표문에서 제시한 대안은 최대 엔트로피 조화문법(Maximum entropy harmonic grammar)으로 이 모델에서는 변이형 사이의 출현을 양적으로 설명하는 것에 문제가 있다고 본다. 따라서 변이형 사이의 절대적 출현 빈도를 다루는 일종의 수치제약문법이라고 볼 수 있는데, 최대 엔트로피 문법은 정보이론(Shannon 1948)에서 시스템 정보의 양을 나타내는 엔트로피라는 개념을 사용하여 구축한 모델로 변이형의 출현 확률을 조화치(harmony score)에서 직접 얻을 수 있다. 또한 제약에 수치를 부여하여 가중치(weight)를 정하여 변이현상을 설명하는 방식이다. 발표문에서 분석한 영어 어말 /t, d/ 탈락현상은 수의적이다. 탈락이 되는 환경은 휴지 앞, 모음 앞, 자음 앞에서이며 물론 수의적으로 탈락이 적용되지 않을 수 있다. 문제는 규칙기반 문법과 최적성이론과 같은 제약기반 문법의 해결 방식이 범주적이라는 지점에 있다. 따라서 다양한 변이형이 관찰되는 비범주적 음운현상을 해결하려 할 때 발표문의 제안처럼 양적 양상(quantitative pattern)이나 통계적인 문법 모형을 고려하지

않을 수 없다. 발표문에 제시된 영어의 어말 /t, d/ 현상의 분석에 있어서 최대 엔트로피 조화문법을 적용한다면 해당 분절음이 놓이는 환경이나 방언에 따른 차이에 의해 결정되는 변이형에 제약의 강도를 숫자 가중치(weight)로 표시하여 표면형태가 위배하는 제약의 가중치 및 그 위배 횟수에 따라 그 형태의 발생 확률이 부여된다. 따라서 이러한 설명 방식은 동점위배로 변이현상을 설명하는 순위제약문법에서 다루는 변이현상보다 더 설명적 타당성이 있으며 변이형의 출현 확률을 조화치에서 직접 얻을 수 있는 장점이 있다. 즉 영어의 어말 /t, d/ 탈락현상이 주어진 환경에 따라서 삭제 비율이 다르기 때문에 문법 분석에서는 이들의 발생비율을 포착할 필요성이 있다는 것이다. 흥미로운 점은 삭제비율이 환경의 차이뿐만 아니라 방언에 따른 차이점을 보인다는 것이다.

그러므로 문법에서 발생비율을 포착할 수 있다면 환경에 대한 삭제비율뿐 아니라 방언에 의한 발생빈도의 차이를 파악할 수 있기 때문에 과거 규칙이나 제약만으로 해결하려 했던 변이형에 대한 논의를 수리적으로 해석할 수 있다. 결국 경험적인 자료의 기술에서 이론적인 문법분석의 시기에 이르러 음운론은 다양한 시도를 거듭해왔고, 현재는 발표지의 논의처럼 기존의 OT 제약문법에 가중치를 포함하여 정확한 가중치 계산을 위해 문법학습 프로그램을 사용하며, 이를 통해 화자의 발화자료를 학습자료로 삼아 문법학습을 수행함으로써 인간의 문법학습을 모델링할 수 있는 장점이 있다. 따라서 최대 엔트로피 조화문법은 기존 이론에서 해결하기 어려웠던 비범주적 현상을 연구 대상으로 확대할 수 있다는 이론적 장점이 있다. 발표문에서는 영어의 어말 t/d 삭제현상을 예로 들어 설명했지만, 최대 엔트로피 조화문법을 통해 한국어의 음운자질, 삽입모음, 음운의 분포 등을 정보이론의 개념을 통해 음운적 복잡성과 연관을 지어 연구할 수 있을 것이며, 정보이론의 엔트로피와 엔트로피 기여도가 자질의 변별성과 한국어 말소리의 분포 연구에도 역시 적용될 수 있다.

정리하자면 발표문의 논지는 관여하는 소리의 엔트로피는 분포 환경에 따라 얻은 엔트로피의 환경에 따른 확률을 가중하여 구하기 때문에 이러한 논의가 기존의 범주론적 음운론의 해석 방식을 탈피하고 언어 내적, 혹은 방언형에서 나타나는 변이형을 가중치와 확률로 평가할 수 있기 때문에 음운론에서 범주적으로 설명할 수 없는 여러 변이형의 분석에 용이한 이론이며, 나아가 국어음운론에 나타나는 다양한 변이형의 연구에서도 충분히 적용하여 후속 연구가 가능할 것이다.

참고 문헌

- 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.
- Hammond, Michael (1994) Gradience, phonotactics, and the lexicon in English phonology. *International Journal of English Studies* 4.2, 1-24.
- Hayes, Bruce & Colin Wilson (2008) A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39, 379-440.
- Prince, Alan & Paul Smolensky (1993) *Optimality theory: Constraint interaction in generative grammar*. Oxford, UK: Blackwell.
- Shannon, Claude (1948) A mathematical theory of communication, *The Bell System Technical Journal* 27, 379-423.