

## 『훈민정음해례본』에 나타난 성리학 원칙을 중심으로 한 한글 자동 정보 처리

### 문제 제기:

한국의 고유 문자인 한글은 로마 문자와 마찬가지로 음소 문자이지만 단순히 한 줄에 나열하는 일차원적인 로마자와는 달리 글자를 모아쓰는 <음절자적인 다발> (syllabographischer Cluster) 로써 이차원적으로 기재한다는 특징을 지닌다. (음절자적인 다발이라는 개념은 그 사용의 합리성이 자명함에도 불구하고 학술 문헌에서 아직 일반화되어 있지 않다.) 유니코드 2.0 의 도입 이후, 이 문자 코드 규약에 의해 구현이 가능한 음절자의 개수는 약 11,000 에 지나지 않지만, 28 개의 한글 자모를 가지고 실제 손으로 쓸 수 있는 음절자적인 다발의 개수는 이론적으로 약 390 억 개에 이른다. 문자 표현기능에서 드러나는 이러한 컴퓨터상의 제한성은 분석 처리의 한정성으로 인해 한층 더 부각되어 나타난다. 자동 정보 처리에서의 전반적인 장애물은, 그 최소 단위가 한 글자가 아니라 외관상 글자 단위의 조합으로 구성되어 있는 것처럼 보여지지만 실상은 이미 만들어져 완성돼 있는 이른바 <가짜 다발> 이라는 점에서 확연히 드러난다.

### 현재까지의 연구 현황:

한글 문자에 관한 학문적 연구는 지금까지 각기 다른 두 진영에 의해 이루어졌다. 그 하나는 1446 년에 발행된 한글 문자의 기초가 되는 『훈민정음해례본』 및, 이 문헌에 쓰여져 있는 성리학, 즉 도교에 기초한 철학적·우주적 원리를 연구한다. 두번째 진영은 기존 한글 문자 연구의 학문적·이론적 기초를 전혀 염두에 두지 않고, 정보학적 시각에서만 한글 문자 연구에 접근한다.

### 본 논문의 주장 :

한글 문자와 위에 간략히 언급한 자동 정보 처리상 문제를 푸는 열쇠는 성리학의 철학적·우주론적 원리에서 찾아진다. 다시 말하자면, 이 문제는 훈민정음 창제시에 근간이 되었던 공리(公理)적인 시스템을 고려해야만 해결된다고 할 수 있다.

### 방법론적 접근 :

우선 예비 단계로 아직도 종종 제기되고 있는 역사적 자료로서의 『훈민정음해례본』의 가치에 대한 의문을 종식시켰다. 오류없는 접근 방식의 또 하나의 중요한 전제 조건은 문자(Schrift), 글자(Buchstabe), 음소(Phonem), 자소

(Graphem), 음절 (Silbe), 음절자(Syllabogramm), 부호화(Codierung) 등의 핵심 개념에 대한 정리 작업을 하는데 있다.

본 연구 조사의 출발점은, 페터 아이젠베르크 (Peter Eisenberg)에 의해 개발된 문자 구조를 이루는 총체적 모델을 도입하여 이를 한글 문자에 적용한다는 데 있다. 한글의 철학적·우주론적 원칙은 아이젠베르크 모델의 소위 메타차원에 상응하는데, 이는 문자를 형성하는 다른 모든 차원 즉 자소, 자소와 음소의 상호 연관 관계, 그리고 형태소의 차원을 전부 포괄한다.

아이젠베르크의 모델을 이용하면 여러 차원에서의 문자에 관한 규칙성이 선명히 분리되어 드러난다. 그리하여 자동 정보 처리상에 드러나는 한글의 문제점이 이러한 각각의 차원을 고려하지 않고, 서로 다른 차원의 상이한 법칙을 혼합함에서 기인한다는 것이 자명해진다. 한글의 부호화가 그에 적합하지 않는 차원에 설정됨으로 인해 위에 언급한 제한성이 발생한다.

다음의 방법론적 단계는 성리학의 철학적·우주론적 원리, 즉 메타차원에서의 기능적 골격을 제시하는데 있다. 본 연구의 원(原)자료로는 『훈민정음해례본』 및 중국 성리학의 기초이론서를 사용했다. 『훈민정음해례본』의 곳곳에 보여지는 이 중국 문헌에서 직접 인용된 부분들은 훈민정음이 성리학적 원칙에 기초한다는 것을 명백하게 보여주는 것으로 평가할 수 있다.

이 단계에서의 첫번째 구체적인 산물은 한글 자모를 부호화하기 위한 기반으로써, 자동 정보 처리에 기초적으로 중요한, 이론(異論)의 여지가 없는 (biunique) 한글 자모 순서의 확실한 인식이다.

이러한 첫번째 방법론적 단계의 결과를 바탕으로 날자들을 음절자적인 다발로 형성되는 차원 (syllabotaktische Ebene)에서 이차원성, 구조적 불변성 (자음자+모음자+자음자) 및 클러스터링의 자유라는 세가지 중요한 한글의 구조적 속성을 연구하였다.

이 트로이카는 본 논문의 상당 부분을 이끄는 길잡이로서, 한글 문자의 법칙성을 자형학, 입력, 부호화 및 부분적으로 출력 분야에 적용하는데 있어 핵심적 역할을 했다.

본 요약의 한정된 지면상으로 모든 단계를 기술하기가 불가능하지만 결론적으로 확실한 것은 컴퓨터라는 도구에서도 한글 문자 390 억 개의 음절

자적인 다발의 전체를 28 자모로만 구성할 수 있다는 것이다. 그리고 여기서 전제조건은 문자를 구성하는 각각의 차원에서 전반의 구조적 법칙이 고려되어야 한다는 것이다. 그리하면 말뭉치나 다른 언어자료의 학문적 분석에 있어서 더 이상의 제약이 없어진다고 해도 과언이 아니다.