

훈민정음 해례 기반 KS X 1026-1:2007의 평가

권기홍, 변정용
동국대학교 과학기술대학 컴퓨터공학부
{yano3000,byunjy}@dongguk.ac.kr

An Evaluation of KS X 1026-1:2007 Based on Huminjeongeum Principle

Ki-Hong Kwon, JeongYong Byun
School of Computer Engineering, Dongguk University at Gyeongju

요 약

1987년 완성음절형 한글 부호계(KS C 5601-1987)가 처음 제정된 이래 1992년 ISO 10646-BMP 또는 Unicode에 정보교환용 한글부호계를 반영하기 시작하여 2005년에 KS X 1005로 국제문자부호계를 제정하고 2007년 최종적으로 KS X 1026-1 제정을 통하여 정보교환용 한글 처리지침을 마련하는 것으로 옛한글처리를 완성하였다고 한다. 본 논문에서 훈민정음창제원리에 입각하여 옛한글처리 완성도에 대하여 정보교환용 한글처리의 한계와 문제점을 평가함으로써 한글처리의 효율성을 기하고자 한다.

1. 서론

행정전산망이 추진되면서 1987년 완성음절형 한글 부호계(KS C 5601-1987)가 처음 제정되었다. 그 후 1992년 국제문자호계 기본다국어판(ISO/IEC 10646-BMP) 또는 유니코드(Unicode)에 정보교환용 한글부호계를 반영하기 시작하여 자모형[1], 자소형(grapheme), 완성음절형 등 3개의 부호계를 반영하였다. 이를 2005년에 KS X 1005 국제문자부호계를 제정하였고, KS X 1026-1:2007[2]를 제정하여 정보교환용 한글 처리지침을 마련하는 것으로 매듭짓고 있다. 이것으로 컴퓨터에서 한글처리가 완결된 것처럼 말하는 것은 속단이다. 본 논문에서는 훈민정음창제원리에 따라서 KS X 1026-1: 2007의 한글처리지침이 처리하는 방식을 평가하여 한글부호계의 한계와 문제점을 지적하고, 해결 방안을 제시하고자 한다.

2. 기존 연구

1987년 이래 훈민정음 해례에 기반한 한글부호계 연구는 다방면에서 진행되어 왔다[3][4]. 1992년 ISO JTC1/SC2 서울회의에서 훈민정음에 따른 초성,중성,종성 등 자소가 최초로 반영되었다. 하지만 복자소 240자가 등록되어 이후 많은 문제를 낳게 되었다. 2013년엔 정음형 웹입력기 개발[3]등 지속적인 연구가 이루어지고 있다.

3. 훈민정음과 한글부호

3.1 훈민정음의 과학성

1443년 창제된 훈민정음은 제자해, 초성해, 중성해, 종성해를 기본으로 해서 기본 자모 45자를 제정한다. 여기서

‘중성은 초성을 다시 쓴다’고 하여 중성 17자의 표기를 생략함으로써 28자로 축약된다. 그리고 합자해에서 이를 활용하여 천지자연의 소리를 나타낼 수 있도록 4가지 확장법칙을 정의한다. 첫째 입술소리 아래 o를 붙이면 입술가벼운 소리가 된다(연서법). 둘째, 초성, 중성, 종성을 2자, 3자 합용병서한다(병서법). 셋째, 가로꼴 중성은 초성 아래, 세로꼴 중성은 초성 오른쪽에 쓴다. 그리고 중성은 잇대어 적는다(부서법). 넷째, 모든 글자는 중성을 중심으로 모아야 소리를 이룬다(성음법) 등이다. 이를 통하여 약 399억음절을 생성할 수 있다. 이것은 해례 정인지서에서 언급한 “천지자연의 소리가 있으며 천지자연의 문자가 있다”와 배우기 쉽기로는 “지혜로운 이는 아침먹기 전에, 그리고 어리석은 자라도 열흘이면 깨우칠 수 있다”고 한 명제들은 훈민정음이 과학적이라는 증거들이다.

3.2 한글 정의의 문제점

국립국어원의 표준국어대사전에 따르면 ‘우리나라 고유 문자의 이름. 세종 대왕이 우리말을 표기하기 위하여 창제한 훈민정음을 20세기 이후 달리 이르는 것으로, 1446년 반포될 당시에는 28 자모(字母)였지만, 현재는 24 자모만 쓴다.’로 되어 있다. 이것은 사전적 의미이다. 이를 뒷받침한 규정은 국립국어원에서 정한 어문규정 중에 하나인 ‘한글맞춤법’이다. 여기에는 자모와 우리말 표기에 관련된 전반적인 규범이 정의되어 있다. 여기서 중대한 오류는 자모라는 호칭과 자모의 숫자의 문제가 있고, 훈민정음 해례의 합자해 부분을 완전히 무시하고 있다. 한글의 음절은 11172자 인데, 훈민정음은 약399억자로 360만배에 달한다.. 그래서 한글은 훈민정음과 대등하지 않으므로 앞의 사전

에서 ‘훈민정음을 20세기에 달리 이르는 것’이라고 함은 틀린 말이다.

3.3 KS X 2016-1:2007[2]의 문제점

훈민정음은 15세기 음운연구의 수준 높은 결과물이다. 따라서 그에 상응하는 용어 사용은 물론 창제원리에 입각하여 서술하여야 훈민정음의 참가치를 유지시킬 수 있기에 다음과 같이 지적한다.

첫째, 용어정의 3.3의 ‘낱자’라 함은 애매한 표현이다. 왜냐하면 훈민정음에선 소리의 구성요소를 초성, 중성, 종성으로 정의하고 있기에 소리의 구성과 관계없이 낱자라 함은 훈민정음의 용어가 아니며 부호계에서도 큰 의미가 없다. 더군다나 복자음과 복모음도 낱자라고 칭하고 자음과 모음과 같은 의미로 본다는 점 또한 잘 못된 표현이다.

둘째, ‘3.5 현대한글’의 정의는 ‘한글맞춤법[1]에 따라 표현할 수 있는 자소와 소리마디(음절)의 범위이다’로 해야 한다. 그리고 ‘3.6 옛한글’은 ‘훈민정음 창제 이후 사용한 글자 가운데 한글맞춤법에서 제외된 낱자소 또는 합자소를 이른다’로 해야 올바르다.

셋째, ‘4.1 조합한글낱자’는 기존 한글자소가 배정된 영역인 U1100 - U11FF의 빈공간에 초성 5자, 중성 5자, 종성 7자를 배정하고, 다시 한글자소 확장-A UA960 - UA97C에 초성 29자, 확장-B U7DB0 - UD7C6에 중성 23자와 UD7CB - UD7FB에 종성 49자 등 최근 새로 발견된 117자[1]를 배정하였다. 이렇게 산발적으로 배정된 연유는 추가 배정하면서 공간이 빈 곳을 찾아서 배정한 결과이다. 결과적으로 가나다순으로 자연스런 정렬은 할 수 없다는 것이다. 정렬의 문제는 기존 한글자소 영역에서도 현대한글자소를 먼저 배정한 뒤 옛한글자소를 배정하여 정렬 규칙을 기본적으로 파괴하고 있다.

다섯째, 이와 관련된 자료를 보면 옛한글 자소 117자를 추가하여 국제표준에 등록된 한글 자소가 초성 124자, 중성 94자, 종성 137자이며 총 355자가 됨에 따라 현대한글은 완성음절형 11172음절자, 그리고 옛한글을 포함하여 약 160만 음절자를 표현할 수 있다고 한다. 하지만 훈민정음해례는 약 399억 음절자를 표현할 수 있다.

4. 평가

KS X 2016-1:2007의 가장 큰 오류는 한글이 마치 훈민정음과 같은 능력이 있다고 보는 점이다. 한글은 훈민정음일 수 없다. 그럼에도 표준국어대사전에서는 한글은 훈민정음의 다른 호칭이라 한다. 한글은 국립국어원 어문규정의 한글맞춤법에 근거하고 있다. 한글맞춤법은 훈민정음해례의 극히 일부분을 따른다. 그러면 표현범위를 보자.

- 1) 현대한글의 표현범위는 초성 19자 x 중성 21자 x 종성 28자(받침없음 포함) = 11172 음절자
- 2) KS X 2016-1:2007에 따른 표현범위는 초성 124자 x 중성 94자 x 종성 137자 = 약 160만 음절자
- 3) 훈민정음의 표현범위는 399억 음절자

KS X 2016-1:2007은 크게 한글맞춤법을 위반한다. 더 심각한 문제는 한글맞춤법으로 훈민정음해례의 문자표현을 하려는 점이다.

1992년 ISO JTC1/SC2 서울회의에 대비한 국가안 도출을 위하여 국내 JTC1/SC2 회의에서 복자소 240자가 등록되었다. 그래서 2007년에 새로 발견된 117자를 등록할 때 산재되어 배정하게 된 것이다. 이렇게 된 근본 이유는 복자소를 하나의 낱자로 정의하고, 한글맞춤법의 규정에 따랐기 때문이다. 하지만 훈민정음 해례는 합용병서를 통하여 복자소를 생성한다. 훈민정음해례는 합자해에서 4개의 법칙으로 자소 확장을 규정하고 있다. 이 부분이 바로 훈민정음이 과학적인 문자라는 것이다. 수학의 집합에서도 원소나열법과 조건제시법이 있는데 훈민정음은 전체 문자 집합을 원소나열법이 아닌 조건제시법을 사용하여 표현한 것이다. 이것이 ‘지혜로운 이는 아침 먹기 전에 어리석은 이라도 얼흘이면 깨우친다’고 한 것이다. 또한 ‘천지자연의 소리를 적을 수 있는 천지자연의 문자’라는 것이다. 그래서 이것을 훈민정음의 과학성이라고 한다.

5. 결론

훈민정음해례의 음절 생성원리를 바탕으로 해서 한글처리 지침인 KS X 2016-1:2007의 성능을 평가해 보았다. 현대한글은 한글맞춤법에 근거하고 있으며 표현범위는 11172음절자이고, 확장해서 160만자이다. 1992년 이래 117자 추가로 발견되어 한글자소 구역에 산재시켜 추가하였다. 이것은 정렬 알고리즘의 복잡도를 극히 높게 된다. 그런데 앞으로 계속 발견되면 복잡도는 더욱 악화될 것이다. 이러한 결과가 발생하는 이유는 한글맞춤법의 그릇에 훈민정음해례에 기반한 옛한글을 담으려는 시도 때문이다.

훈민정음해례가 생성하는 음절 표현범위는 무려 399억 음절이기 때문에 훈민정음해례의 원리를 따르지 않으면 옛한글을 포함한 한글처리는 매우 복잡해지고 과학적인 처리가 불가능하다. 따라서 훈민정음의 낱자소 45자를 기본으로 하고, 합자해의 4개 법칙에 따라서 399억 음절을 생성하는 방식을 따랐을 때 완전하게 과학적으로 처리할 수 있음을 확인할 수 있다.

참고문헌

- [1] 어문규정-한글맞춤법, 국립국어연구원
- [2] KS X 1026-1:2007 정보기술-국제문자부호계-한글처리지침, 산업자원부, 기술표준원
- [3] 이형준, 변정용, “완전조합 음절지원 훈민정음 웹입력기”, 컴퓨팅의 실제와 레터, 정보과학회논문집, 제19권 제6호, pp371-375, 2013.6
- [4] JY Byun and SH Oh, "A Representation of Korean Syllabic Characters defined in Hunminjeongeum," 13th Int'l Multimedia Modeling Conference, LNCS (Singapore), Proceedings Part II, SPRINGER VERLAG BERLIN, pp.662-670, Jan. 2007