

훈민정음의 창제원리와 한글부호계의 제정원리 연구

변정용* 임해철**

* 동국대학교 전자계산학과 ** 홍익대학교 전자계산학과

A Study on The Principle of Hunminjeongeum Creation And The Principle of Hangul-Code Regulation

Byun Jeongyong* Lim Hae Cheol**

*Donguk Univ. Dept. of Computer Science **Hongik Univ. Dept. of Computer Science

요 약

정보교환용 한글부호계는 훈민정음 창제원리와 ISO 2022 권고안이 그 바탕이 되어야 현행의 모든 한글 부호계가 가진 문제의 해결이 가능하다. 한글 부호계는 정음 28자를 기본 체계로 하여 언서법, 병서법, 부서법, 성음법등에 의한 전환을 통하여 다양한 소리를 표기할 수 있도록 규정하고 있다. 그리고 한글 옛글자를 포함한 모든 한국어 정보처리가 가능하도록 부호계가 제정되어야 하고 두벌식 낱자형에서 자음의 이중성으로 인한 부담을 개선한 세벌식 낱자형 부호계를 제안하고 궁극적으로 정보교환용 한글 부호계의 개선도 꾀한다.

1. 서론

컴퓨터를 이용한 한국어 정보처리는 한국어의 표기체계인 한글문자를 그 속성에 맞는 부호계로서 적합하게 표현하는 것이 그 첫째 과제이다. 여러번에 걸친 한글부호계의 제정과 개정의 결과 현재는 KS C 5601-1987 일명 완성형 부호계[10]를 정보교환용 한글 한자 부호계의 표준으로 삼아 사용하고 있다. 그런데 완성형 부호계에 대한 문제는 1987년 개정 당시 부터 한글 표현에 대한 부적합성이 컴퓨터학과 국어학과 그리고 일반 사용자들로부터 논란의 대상이 되어 왔으며 개정을 하여야 한다는 의견이 지배적이다.

한글 부호계 문제의 근원적 해결은 한글문자 구조의 모태인 훈민정음의 창제원리에서 찾아야 할 것이다. 훈민정음 예의편과 해례[3]에서 한글의 구조를 음양오행 원리에 기반하여 기본 자모와 생성규칙을 규정하여 훈민정음의 정신을 구현하고 있다. 이러한 원리에 반하여 현행의 완성형은 무려 2,350자나 되는 한글 글자를 부호화 대상으로 삼고 있으며, 이는 현대글자 중에 4/5 글자의 표현에 제약가한 것으로[6] 훈민정음의 창제원리에 크게 벗어난다. 또한 처리과정에서 입력 글자판 이후에 완성형 부호가 적용되어 한국어 정보처리의 적용성을 크게 떨어뜨리고 있다.[7] 이의 극복은 한글 부호계 제정 원리에 훈민정음의 창제 원리를 적용하는 것이다.

본 논문은 한글문자의 속성과 부호계의 제정원리를 파악하여 문자의 속성 표현과 부호계 제정 원리를 동시에 수용하면서 정보처리의 효율성을 제고하도록 한다는 원칙을 설정하고, 훈민정음의 창제원리에 정보교환용 부호계 제정에 관한 국제표준 권고안인 ISO 2022의 제정원리[12,13]를 적용하여 현행 한글 부호계가 가지고 있는 한계성을 근본적으로 해결하고 한글 관련 소프트웨어의 생산성과 지적노동생산성을 향상시켜 줄 세벌식 낱자형 한글 부호계로의 개정 방향을 제시하고자 한다.

2. 정보교환용 부호계의 제문제

2.1 정보교환용 부호계의 사용목적

한글 부호계를 올바르게 제정하고 사용하려면 우선 정보교환용이라는 용어의 뜻과 적용범위를 알아야 할 것이다. 이들에 대하여 한국공업규격에서 정의하고 있는 내용은 다음과 같다.[10]

1) 용어의 뜻: 서로 다른 시스템 사이에서 정보를 이용하기 위하여 하나의 시스템에서 다른 시스템으로 정보를 전하는 것

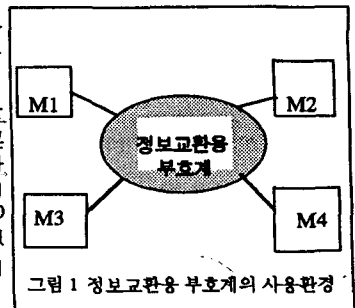
2) 적용범위

o 이 규격에서는 한글, 한자, 특수문자, 외국문자중 일반적인 국어문장에서 표기되는 문자에 대하여 규정한다.

o 정보교환용 부호는 정보교환의 기본이 되는 부호로서 서로 다른 부호계를 사용하는 시스템간의 정보교환시 우선이되는 정보표현 형식임을 의미한다.

o 정보교환은 대상 시스템간에 연결된 회선을 통하여 이뤄지는 온라인 형태와 정보 기록매체에 기록되어 전달되는 오프라인 형태로 구분할 수 있으며 본 부호는 양형태에 모두 적용된다.

상기의 내용에서 보는 바와 같이 정보교환용 부호라 함은 시스템간의 정보교환을 위한 통일적인 부호계를 의미한다. 즉 시스템마다 내부적으로 다른 부호계를 사용하고 있다고 하더라도 통신에서는 정보교환용 부호계를 축으로 하여 상호 전달하도록 하는 목적을 가진 부호계이다. 적용범위에서 일반적인 국어문장에서 사용하는 문자라 함은 완성형에서 통계적 빈도수에 의하여 선정된 한글 글자 2,350자를 말한다. 그러나 국어학계에서는 320 가지 경우의 표현할 수 없는 글자의 사례를 보이고 있다. 그리고 부속서에는 난립하여 있는 16단위 조합형과 7단위 낱자형의 통일을 유도하기 위한 표준 부호계안을 부속서에서 제안하고 있다.



2.2 정보교환용 한글부호계의 계통

한글 부호계의 문제는 옛글자를 포함한 한글의 모든 가능한 글자 집단의 표현, 다양한 특수 문자의 표현, 한자의 표현과 처리 등이 가능한지 그리고 국제적인 통신등에 부합한지 등이 종합적으로 고려 평가되어야 할 것이다. 여기서 기존의 부호계 즉 초성 및 중성의 모든 자음의 집합인 자음 30자와 모음 21자를 표현한 두벌식 낱자형 부호계와 초성, 중성, 종성을 각각 5비트씩 할당하여 16비트로 표현한 조합형 부호계, 2,350 자의 한글과 4,888자의 한자등을 나열하고 일련번호를 부여한 형행 표준안인 완성형 부호계[10]등에서 현재 발생하는 문제들을 들어 본다.

첫째, 정보교환용 부호로서의 적합성을 보자. 컴퓨터상에서 한글처리는 문자열 비교, 정렬, 검색등과 같은 내부처리와 서로 다른 시스템 간의 통신과 같은 외부처리가 있는데 정보교환용 부호는 외부처리에 해당한다. 외부처리에서 중요한 사항은 같은 부호계를 사용하는 시스템간에는 문제가 되지 않지만 부호계가 다를 때 시스템간의 정보 전달시에 부호를 번역해야 하는 데 그럴려면 모든 글자의 표현이 가능하여 그들의 축이 될 수 있어야 한다. 그리고 시스템간의 통신도 원활히 이루어야 한다.

먼저 부호계 번역의 축으로서의 기능을 보면,

가) KS C 5601-1974 낱자형은 한글의 현대글자에서 필요로 하는 모든 자모를 가지고 있으므로 초성, 중성, 종성을 분리하여 표현하면 어떠한 글자라 하더라도 표현이 가능하고

나) KS C 5601-1982 조합형은 2 바이트를 초·중·종성에 각각 5비트씩 할당하므로써 한글 표현에 낱자형과 함께 문제가 없다.

다) KS C 5601-1987 완성형은 통계상 빈도수 0.001%이상의 2,350자의 한글 글자만이 가능한 데 현대 글자의 총 조합수인 11,172자의 1/5에 해당하므로 축으로서 문제가 된다. 물론 2,350자 외의 글자를 송신할 때는 낱자 부호에 111 부호를 앞세워 보내는 방식을 제시하고 있으나 능률적인 처리를 한다는 면에서 이는 예외 사항을 부여하며, 한 글자를 보내려고 8 바이트를 전송하는 부담이 있다. 또한 글자 수 부족에 대한 보완 방안으로 최근에 제공한 보조 글자판이 있으나 그것은 예외사항 처리 부담을 더욱 더 가중시키는 결과를 낳는다.

그리고 통신에서의 문제는 도형문자영역과 제어문자영역을 엄격히 준수하는 것으로

가) 낱자형과 완성형은 도형문자들을 해당 영역에 정확히 표현하였으므로 문제가 없고,

나) 조합형은 중성 5비트중에 하위 3비트에 의하여 일부 제어문자영역을 침범한다. 그리고 한 글자의 표현함에 있어서 두 바이트가 7비트와 8비트의 혼합환경으로 표현된다. 이것은 앞 바이트의 8번째 비트내용이 1이 되어 한글 표현시 앞 바이트는 항상 8비트 환경에 속하게 된다. 그런데 뒷 바이트는 중성의 5비트 중에 3비트가 할당되는 데 맨 왼쪽 1비트는 뒷 바이트의 8번째 비트이다. 이것의 내용은 모음을 배정하였을 때 32중에 16개가 0이 할당되어 표상에서 11번은 7비트 환경에 소속하게 된다. 그리고 제어문자영역의 침범은 3비트중 왼쪽 2비트의 내용이 00 또는 01이 되는 경우로 모음이 배정되었을 때 8개가 있는데 이들이 제어 문자영역을 침범하는 주역이다. 이것이 바로 ISO 2022 권고안에 위배되는 점들로서 조합형에선 피할 수 없다는 부분이다.

둘째, 한자의 표현 문제는 낱자형에서는 고려된 바 없고 조합형에서는 별도의 부호를 앞에 세워서 사용하며 완성형은 4,888자를 제공한다. 이 외에도 KS C 5714-1977은 일본의 JIS C 6226과 같은 한자 부호계이며 특히 KS C 5619-1982는 현재의 완성형과 같은 부호계로서 한자 1692자를 가지고 있다.[10, :1]

셋째, 옛글자의 표현은 낱자형 조합형 모두 없으며 완성형은 옛글자의 낱자를 모두 제공하고 있다. 그러나 글자를 표현하지는 못하고 있는데 1990년 보완 과정에서 약 1,500자의 옛글자를 추가한다고 하였지만 모든 글자집합을 표현하기는 어려울 것이며 이 또한 현대 글자와 더불어 사전식 나열이나 문자열 정렬을 할 때 알고리즘을 만들기 가 어려울 것이다.

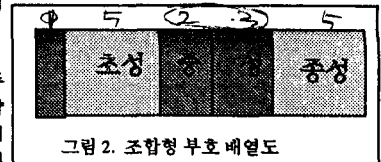


그림 2. 조합형 부호 배열도

3. 훈민정음창제원리

3.1 문자론

문자의 발전단계를 자형의 구성법에 따라서 1단계 상형자형(표어문자), 2단계 회의자형(표어문자), 3단계 형성자형(표어문자), 4단계 수음자형(표음문자), 5단계 표음자형(표음문자)으로 나눌 때[9] 한글문자는 제 4 단계의 표음문자에 속한다. 그리고 일본의 가나는 3단계의 반차자형으로 표음문자이되 음절문자이다. 로마자는 제 4단계의 표음문자이면서 음소문자에 해당한다. 한글문자는 낱자개개는 훈민정음 해례의 제자해에서 보는 바와 같이 발음기관의 모양을 본 떠서 만든 상형자이며 소리를 표현하는 표음문자에 속하며 음소문자와 음절문자의 성질을 동시에 가지고 있다. 이는 같은 제 4 단계에 속할 지라도 로마자와는 달리 이차원[8]로 표기함에 기인한다.

3.2 훈민정음 구조원리

훈민정음 구조원리의 역학적 고찰[14]은 훈민정음 해례에 나타나 있는 바와 같이 음운학과 성리학을 앞대배경으로 하여 이루어졌는데 여기서 역학의 원리론, 사시지리, 삼재지리, 순환지리가 모두 적용되었으며, 그리고 음양오행과 상형설, 온학은 제자해의 직접적인 일관된 이론이라고 한다.[3]

정음 28자의 기원은 "상형이자방고호"라 하는데 초성 기본음 ㄱ ㄴ ㄷ ㅇ 등은 아설순치후의 발음기관의 모양을 본 떠서 상형자이고 소리의 거셀과 발음형태에 따라 기본형의 위 또는 아래에 획을 더하여 자음이 확정되어 자음은 17자가 된다. 중성 기본음 ㅡ ㅣ 등은 천지인의 모양을 본 떠서 만든 것이며 천지 또는 천인이 초고 또는 재고의 뜻에 따라, 와 ..을 ㅡ의 위 아래 ㅣ의 좌우에 가획하여 만들면 11자가 되고 중성은 초성과 같다고 하여 28자를 이룬다. 28자는 기본 체계요 그 나머지는 합용과 병서에 의하여 여러가지 소리모양을 표기하도록 되어 있다. 이른바 자음에서 같은 글자를 합한 것을 각자병서라 하고, 다른 글자를 2자, 3자 합해서 쓰는 것을 합용병서라 한다. 그리고 모음에서 2자, 3자를 합한 것을 합용이라고 해례에서 구별하고 있다.[3]

여기서 우리의 관심은 자음 17자와 모음 11자를 연서와 병서를 통하여 자음과 모음을 확장하는 것이며 이는 훈민정음 예의편에 다음과 같이 명시하고 있다.

가) ㅇ 을 순음 아래에 이어 쓰면 입술 가변운 소리되느니라 (연서법)

나) 초성자를 어울려 쓰면 나란히 쓰라 (병서법)

기존의 부호계에서 훈민정음의 원리에 가장 충실한 낱자형 KS C 5601-1987을 보완한 51자 부호계의 제안이 있었고 이 부호계도 이미언서법과 병서법이 적용된 경우여서 자소의 문자집합의 크기가 매우 크다. 그래서 글자판의 33자 부호계를 제안[6]이 있었는데 병서법의 적용하여 생성된 자모는 초성 5자와 중성 4자가 그것이다. 이들은 거의 이미 단음화된 경우여서 현대 글자에서는 기본 자모라고 할 수 있다. 그런데 한글 옛글자의 표현이나 내부처리 성능의 개선등과 같은 요구도에 대하여 만족할 수가 없어 보다 확장된 부호계를 제안한다. 한글 부호계 제정을 위한 요구도는 다음과 같다.

가) 훈민정음 창제원리와 정신에 부합: 기본자모 28자 + (언서법, 병서법, 부서법, 성음법), "무슨소리 무슨말이라도 표현할 수 있다", "전환이 무궁하고 간단하며 요긴하고 정밀하여 잘 통한다", "배우기 쉽고 쓰기 편하다"

나) 한글의 모든 응용분야에 부합: 옛글자 + 현대글자에 대한 모든 한글 및 한국어 정보처리 응용분야

다) 국제표준기구의 ISO 2022의 권고안 적용

라) 내부처리 성능의 개선: 기본자모의 초성, 중성, 종성을 세벌식으로 표현, 훈민정음에서도 구분

리)항은 두벌식 낱자형에서 내부처리시에 자음이 초성과 중성을 한 벌로 하고 있어서 실행중에 자음의 자질을 알아내야 하는 부담을 개선하는 요구이다. 이에 대하여 단모음화된 모음이나 빈도수가 높은 쌍자음을 제외하고는 언서법, 병서법의 적용을 확연히 표시 때로 지연한다. 그리고 한글의 옛글자를 현대글자와 같은 율타리안의 적용분야로서 설정하여 동시에 표현이 가능하게 하였다. 그래서 자음과 모음의 집합의 크기를 최소화 하였다. 이렇게 해서 훈민정음이 초성, 중성, 종성을 구별함을 본 받아 초성과 중성의 자음 두벌과 중성 모음 한벌을 합한 세벌로 기본 자모 체계를 표현하고 나머지는 규칙으로 표현하여 한글 글자의 모든 가능한 경우를 정의하고 실용적으로 쓸 수 있도록 하였는데 그림 5에서 보는 바와 같다.

	8	9	A	B	C	D	E	F	
0				ㄱ	ㄴ	ㅇ	·	ㅏ	ㅑ
1				ㅋ	ㆁ	ㅅ	ㅓ	ㅕ	ㅗ
2				ㆁ	ㅆ	ㅈ	ㅗ	ㅛ	ㅜ
3				ㄴ	ㅈ	ㅊ	ㅓ	ㅕ	ㅛ
4				ㄴ	ㅈ	ㅊ	ㅓ	ㅕ	ㅛ
5				ㄴ	ㅈ	ㅊ	ㅓ	ㅕ	ㅛ
6				ㄴ	ㅈ	ㅊ	ㅓ	ㅕ	ㅛ
7				ㄴ	ㅈ	ㅊ	ㅓ	ㅕ	ㅛ
8				ㄴ	ㅈ	ㅊ	ㅓ	ㅕ	ㅛ
9				ㄴ	ㅈ	ㅊ	ㅓ	ㅕ	ㅛ
a				ㅏ	ㅑ	ㅓ	ㅕ	ㅛ	ㅜ
b				ㅏ	ㅑ	ㅓ	ㅕ	ㅛ	ㅜ
c				ㅏ	ㅑ	ㅓ	ㅕ	ㅛ	ㅜ
d				ㅏ	ㅑ	ㅓ	ㅕ	ㅛ	ㅜ
e				ㅏ	ㅑ	ㅓ	ㅕ	ㅛ	ㅜ
f				ㅏ	ㅑ	ㅓ	ㅕ	ㅛ	ㅜ

그림 5. 세벌식 한글 낱자형 부호계

5. 결론

여러번에 걸친 한글 부호계 개정에 있어서 제정자들의 관점은 한글문자 구조원리의 모태인 훈민정음을 바탕하지 않으므로 인하여 한글문자를 로마자의 속성에 맞추려거나 또는 가나에 맞추려고 하였다. 이러한 노력은 결국 한글의 본성을 상실하게 하고 그러므로써 가장 최근에는 한글이 음절문자 특성만을 가진 것으로 취급하여 독창적이고 과학적이라는 한글이 현대 과학의 산물인 컴퓨터를 맞서는 가장 형편없는 문자로 전락한 완연형이라는 해괴한 문자로 둔갑하여 버린 것이다. 이러한 본질적인 접근은 접어 둔채 그것을 보완한 결과로 한글은 누더기가 되었고 컴퓨터에서 만든 사실상 훈민정음 이진으로 퇴보 하고 만 것이다.

결론적으로 이제까지 한글 부호계의 제정과 개정 과정에서 빛은 혼란의 근원은 훈민정음의 창제 원리를 충분히 반영하지 않음과 문자본적인 접근의 꺾임에 비롯하였다고 말할 수 있다. 따라서 한글 부호계의 제정은 그 기본 바탕을 훈민정음 창제원리에 두고, 문자본적으로 표음문자로서 음소문자와 음절문자의 특성을 함께 가지고 있음을 고려하여 국제표준기구의 권고안을 적용하여 제이의 훈민정음을 창제한다는 각오로 개정하여 나가야 한다고 주장한다. 그렇게 하면 한글 및 한국어 정보처리의 모든 응용분야를 만족시킬 것이다. 그것은 훈민정음이 한글의 모든 낱자 정보를 기본 자소로 표현하고 나머지 복자모와 글자의 생성은 언서법, 병서법, 부서법, 성음법등의 글자 생성규칙으로 표현하여 한글 글자의 모든 집합을 정의하고 있기 때문이다.

앞으로 입력 글자판의 벌식 및 자판의 글쇠수에 관한 좀 더 깊은 연구가 따라야 할 것이다.

참고문헌

- [1] 고영근, 중세표준국어문법론, 탑출판사, 1987
- [2] 남광우, 보중 고어사전, 일조각, 1987
- [3] 서병국, 신강 훈민정음, 학문사, 1978
- [4] 전태규, 이두 고어 사전, 삼양출판사, 1981
- [5] 허용, 옛달본, 과학사, 1985
- [6] 변정용, 한글코드의 개선과 알고리즘의 개발: 낱자형 33자 코드, 동국대학교 경주캠퍼스 논문집, 제 8집, 1989
- [7] 변정용, 한국어정보처리용 한글부호계의 적합성 평가, 동국대학교 경주 캠퍼스 논문집, 제 9집, 1990
- [8] 정희성, 한글문자의 구조와 구성원리에 관한 과학적 고찰, 전자통신, 제 10 권 제4호, 1989
- [9] 니시다, 세계의 문자, 대수관서점, 1981

- [10] 공업진흥청, KS C 5601-1987, 1982, 1974, KS C 5619-1979 KS C 5714
- [11] 일본공업표준조사회, JIS C 6220, JIS C 6226, JIS C 6228, JIS C 6225
- [12] Jim Flemming and William Frezza, "NAPLPS : A New Standard for Text and Graphics", BYTE, 1983
- [13] 송본언길, 일본어 표현용 프로토콜 시안 1,2,3, 인터페이스, 1986
- [14] 이정호, 훈민정음의 구조원리: 그 역사적 연구, 아세아문화사, 1978