

# 훈민정음 과학성에 기반한

## 한글부호 표준화

동국대학교 컴퓨터멀티미디어학부  
변 정 용

단기 4338년 6월 13일

동국대-변정용

1

### 차례

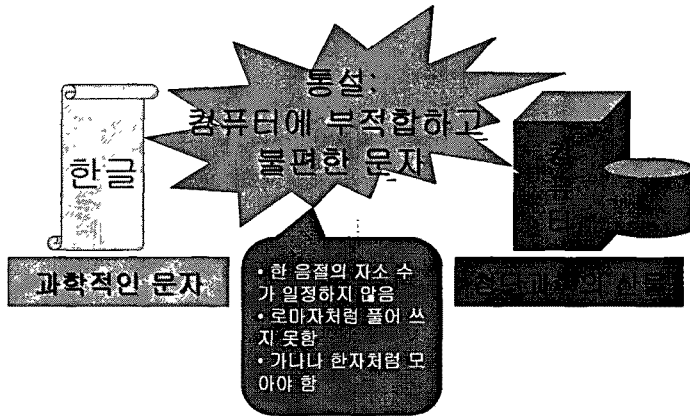
- 서론
- 훈민정음의 과학성
- 한글과학과 표준화 방안
- 구현
- 결론

단기 4338년 6월 13일

동국대-변정용

2

# 서론



단기 4338년 6월 13일

동국대-변정용

# 서론(계속)

## □ 한글부호 논쟁

■ 1987년 전후 조합형이나? 완성형이나?

■ ㈜한글과 컴퓨터의 좌절과 성공

□ 일본의 이찌따로의 패망

□ 세계 유일의 자국 WP

□ 좌절과 815판의 성공

■ MS의 횡포

□ 한컴 hwp와 경쟁과정에서 통합형 한글 부호로 임시 방편

□ MS 창에서 ISO 10646은 짝퉁

## □ 훈민정음 창제원리의 무지

단기 4338년 6월 13일

동국대-변정용

## 훈민정음의 과학성

- 훈민정음 창제원리
- 집합론적 해석
- 문자생성 규칙
- 문자론적 해석
- 컴퓨터와 친화력

단기 4338년 6월 13일

동국대-변정음

5

## 훈민정음 창제원리-1

- 해례본 구성
  - 제자해
    - 발음기관 상형: 아설순치후(ㄱㄴㅇㅅㅇ)
    - 소리의 형성: 천지인(· — |)
  - 초성해
  - 중성해
  - 종성해
  - 합자해
  - 용자례
  - 정인지 서문

단기 4338년 6월 13일

동국대-변정음

6

### 훈민정음 창제원리-2



#### □ 세종의 예견

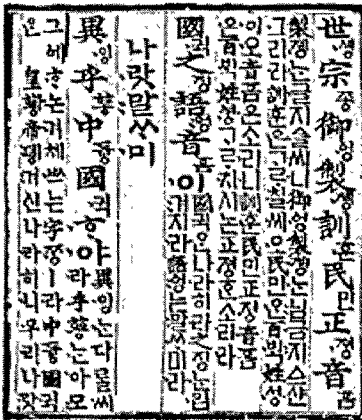
- 500년 후에 컴퓨터가 출현할 것이니, 내 이를 위하여 새로 스물 여덟 자(한글)를 맹가노니...

단기 4338년 6월 13일

동국대-변정용

7

### 훈민정음 창제원리-3



- 훈민정음 예의 본
- 훈민정음 해례 본
- 훈민정음 약사
  - 1443년 창제, 겨울
  - 1446년 해례, 반포
  - 1447년 예의 본 언해

단기 4338년 6월 13일

동국대-변정용

8

## 정인지 서문 -1

- ㄷ 爲螢. ㅁ, 如 ㅅ 爲薪, ㅅ 爲蹄. ㅁ, 如
- : ㅂ 爲虎, : ㅅ 爲泉. ㅅ, 如 : ㅅ 爲海松, ㅁ
- 爲池. ㅁ, 如 : ㅅ 爲月, : ㅅ 爲星之類

### □ 정인지 서문

- 有天地自然之聲, 則必有天地
- 自然之文, 所以古人因聲制字
- 以通萬物之情,

단기 4338년 6월 13일

동국대-변정용

9

## 정인지 서문 - 2

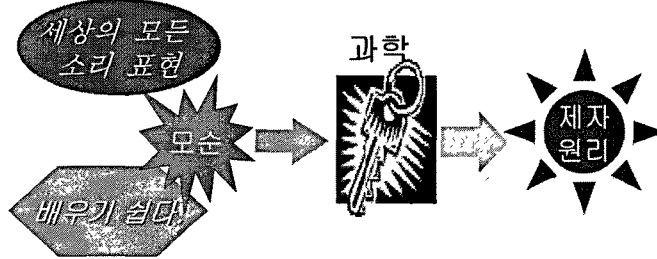
- 括. 以二十八字而轉換無窮, 簡
- 而要, 精而通. 故智者不終朝而
- 會, 愚者可浹旬而學. 以是解書,
- 可以知其義. 以是聽訟, 可以得
- 其情. 字韻則清濁之能辨, 樂歌
- 則律呂之克諧. 無所用而不備,
- 無所往而不達. 雖風聲鶴戾, 鷄
- 鳴狗吠, 皆可得而書矣. 遂

단기 4338년 6월 13일

동국대-변정용

10

### 훈민정음 창제원리-4



- 有天地自然之聲 卽必有天地自然之文(정인지 서)
  - 천지자연의 소리가 있으면 곧 천지자연의 문자가 반드시 있다.
- 以二十八字而轉換無窮,簡而要,精而通.故智者不終朝而會,愚者可浹旬而學. (정인지 서)
  - 전환무궁하고, 간단요긴하여 잘 통하며, 슬기로운 이는 아침 먹기 전에 깨우치고, 어리석은 자라도 열흘이면 깨우칠 수 있다.

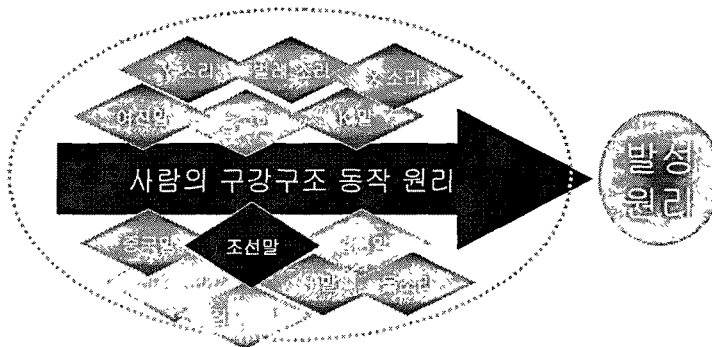
단기 4338년 6월 13일

동국대-변정용

11

### 훈민정음 창제원리-5

- 과학 => 제자원리
  - 어떤 대상에서 발생하는 모든 경우를 꿰뚫을 수 있는 하나의 원리



단기 4338년 6월 13일

동국대-변정용

12

## 훈민정음 창제원리-6 (해례본)

- 제자해 :
  - 天地人 (· 一 丨), 牙舌脣齒喉(ㄱㄴㄹㅅㅇ)
- 초성해 : 17자(ㄱㄴ ... ㅎ △ ㅇ ㅎ)
- 중성해 : 11자(ㅏ ㅑ ... ㅡ ㅣ ·)
- 종성해 : 終聲不用初聲(동일원소 통합)
- 합자해 : 字素(grapheme) 확장
  - 연서법 : 脣音(ㅁㅂㅍ) => 脣輕音(ㄹㄴㅇㅇ)
  - 병서법 : (2, 3字) 各字, 合用
  - 부서법 : 모음위치(가로-아래, 세로-오른쪽)
  - 성음법 : 凡字必合以成音(음절 구성 사유)

단기 4338년 6월 13일

동국대-변정용

13

## 훈민정음 창제원리-7

- 천지자연의 소리 범위
  - 2,3자 합용병서 : 각자병서, 합용병서
    - 초성 :  $17 + 17\Pi_2 + 17\Pi_3 = 5219$
    - 중성 :  $11 + 11\Pi_2 + 11\Pi_3 = 1463$
    - 종성 :  $(17+1) + 17\Pi_2 + 17\Pi_3 = 5220$
  - 조합 가능한 음절 수
    - 초성 5219 x 중성 1463 x 종성 5220
    - = 39,856,772,340 (약 399억)

단기 4338년 6월 13일

동국대-변정용

14

### 집합론적 해석

□ 한글 전체집합 : |H| (조건제시법)

음절자 생성:  
부서법, 성음법

$$|H| = |C1| \times |V| \times |C2|$$

$$= 39,856,772,340$$

**자소확장:**  
2,3자 합용병서

|C1|:5219, |V|:1463, |C2|:5220

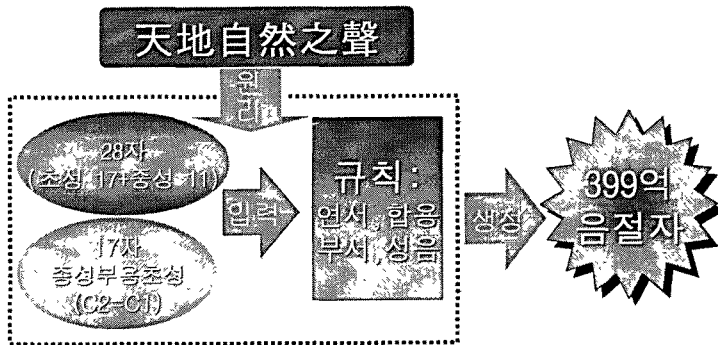
**기본자소:**  
(17자+11자+17자)  
17자(중성부음조성)  
= 28자

단기 4338년 6월 13일

동국대-변정용

15

### 문자생성규칙-1



단기 4338년 6월 13일

동국대-변정용

16



## 문자 생성규칙-2

有天地自然之聲 即必有天地自然之文

28 + 17자

초5219, 중1463, 종5220

부서법, 성음법

약 399억 음절자

훈민정음 문자생성원리도

단기 4338년 6월 13일

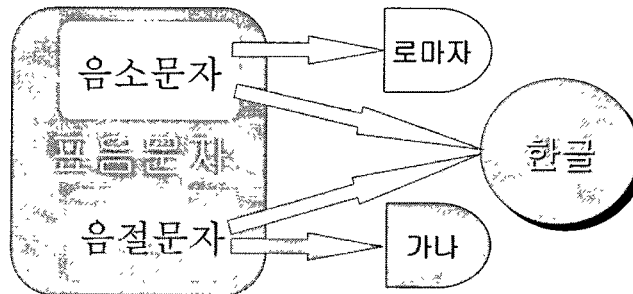
동국대-변정용

17

## 문자론적 해석-1

### □ 표음문자

- 음소문자: 로마자, 음절문자: 가나
- 음소 및 음절 문자 : 한글



단기 4338년 6월 13일

동국대-변정용

18

## 문자론적 해석-2

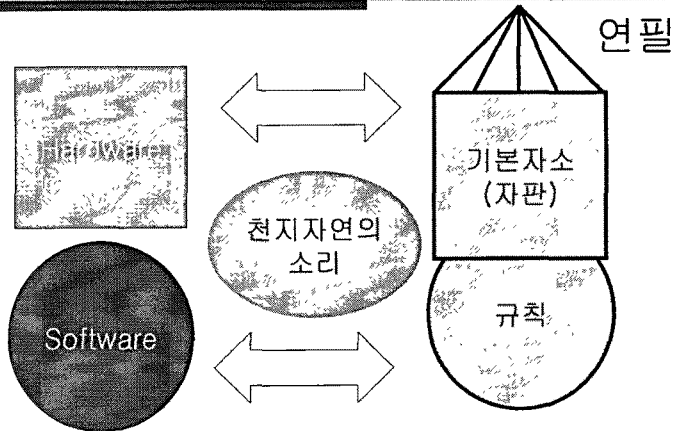
### □ 언어정보

- 언어현상 처리
- 음소 => 기본자소
- 음절 => 기본자소+음절구성규칙

### □ 문자구조정보

- 자소: 초성자, 중성자, 종성자
- 꼴 : 가로꼴, 세로꼴, 양성, 음성, 복자모

## 컴퓨터와 친화력



## 한글 과학과 표준화 방안

- 기존 한글코드의 문제점
- 새로운 요구
- 유니코드의 문제점
- 정음형 코드

단기 4338년 6월 13일

동국대-변정용

21

## 기존 한글 부호의 문제점

- 정보교환용 한글 부호계
  - KS C 5601-1987=> KS X 1001
- 훈민정음 창제 원리 무시
  - 자소집합의 농도 축소
    - 1933년 조선어학회 “한글 맞춤법 통일안” 기준
    - 초성 19 x 중성 21 x 종성 (27+1) = 11172자
  - 연서법, 합용병서법, 부서법, 성음법 없음
  - 옛한글 자소 4자 제외로 옛한글 처리 불가능
- 언어현상 처리를 위한 정보 배제
- 현행 한글표현도 어려움=> 320자

단기 4338년 6월 13일

동국대-변정용

22

## 새로운 요구-1

- 국어정보처리응용의 발전
  - 문자정보처리 => 언어정보처리 중심
  - 개별 응용분야의 통합
- 국제정보통신의 증가
  - 다국어정보처리: 기계번역, 통역
  - 도서, 전자상거래
- 조선글 부호화 표준
  - 북한의 정보산업 육성 정책
  - KPS 9566-1997

단기 4338년 6월 13일

동국대-변정용

23

## 새로운 요구-2

- 남북 자모순 차이=> 가나다 순서 차이
  - 자형순(한글)
  - 음가순(조선글)
    - 초성: ㄱ ㄴ ㄷ ㄹ ㅁ ㅂ ㅅ ㅈ ㅊ ㅋ ㅌ ㅍ ㅎ
    - 중성: ㅏ ㅑ ㅓ ㅕ ㅗ ㅛ ㅜ ㅠ ㅡ ㅣ ㅖ ㅗ ㅛ ㅜ ㅠ ㅡ ㅣ ㅖ
    - 종성: ㄱ ㄴ ... ㅅ ㅇ ㅈ ... ㅎ ㄷ ㅌ ㅍ ㅈ ㅊ ...
- 1996 연길회의 남북자모순 통일안
  - 기본자소 중심 통일안, 가나다 순 통일 불가
- 남북 통합 문자부호 및 처리 방법 개발
  - 가나다순, 전자우편

단기 4338년 6월 13일

동국대-변정용

24

## Unicode / ISO 10646-BMP

### □ 국제문자부호계

- KS C 5700-1995=> KS X 1005
- Unicode를 국내 표준으로 수용
- 국어정보처리는 가능한가? => 자소정보 필요
- 한글부호 전시장
  - Hangeul Jamo(자소형) 240자
  - Hangeul compatibility Jamo(자모형) 51자+x
  - Hangeul(완성음절형) 2350자 + ...(현재 취소)
  - Hangeul(완성음절형) 11172자

단기 4338년 6월 13일

동국대-변정용

25

## Unicode에서 문제점

- KS X 1005-1 : 3 종류의 한글 부호계
  - 완성 음절형, 자모형, 자소형
- 11172자 완성 음절형이 기본 부호계
  - 자연어 처리 및 옛글 표현의 어려움
- 넓은 공간 점유
  - 11172자는 240자 자소형의 부분 집합
- 240자 자소형의 음절 표현의 불완전성
  - 훈민정음 원리의 부분 적용

단기 4338년 6월 13일

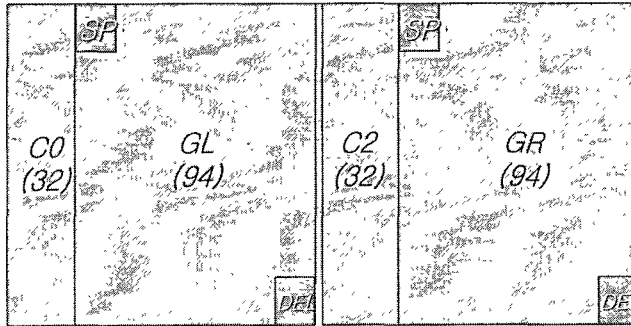
동국대-변정용

26

# 국제문자 코드 규격

□ ISO 646 : 8-bit 환경 (ASCII-8)

- DEL은 7F; 천공카드에서 유래



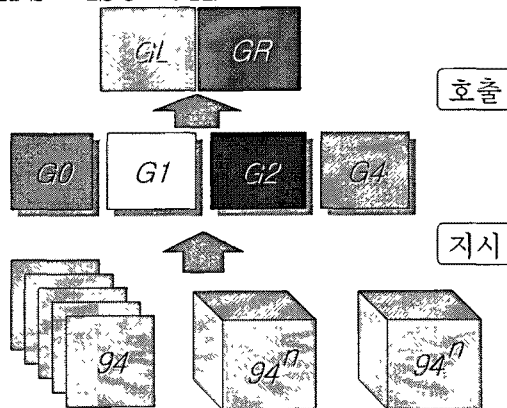
단기 4338년 6월 13일

동국대-변정용

27

# 코드 확장법: ISO 2022

□ NAPLPS=>ISO 2022



단기 4338년 6월 13일

동국대-변정용

28

## 정음형-1996

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
<input type="checkbox"/> 8																
<input type="checkbox"/> 9																
<input type="checkbox"/> A																
<input type="checkbox"/> B																
<input type="checkbox"/> C	F	ㄱ	ㄴ	ㄷ	ㄹ	ㅁ	ㅂ	ㅅ	ㅇ	ㅈ	ㅊ	ㅋ	ㅌ	ㅍ	ㅎ	
<input type="checkbox"/> D	F	ㅏ	ㅑ	ㅓ	ㅕ	ㅗ	ㅛ	ㅜ	ㅠ	ㅡ	ㅣ					
<input type="checkbox"/> E	F	ㄱ	ㄴ	ㄷ	ㄹ	ㅁ	ㅂ	ㅅ	ㅇ	ㅈ	ㅊ	ㅋ	ㅌ	ㅍ	ㅎ	
<input type="checkbox"/> F	F	ㄱ	ㄴ	ㄷ	ㄹ	ㅁ	ㅂ	ㅅ	ㅇ	ㅈ	ㅊ	ㅋ	ㅌ	ㅍ	ㅎ	

## 국제표준규격 ISO 646 개정 제안

### 문자집합의 크기

- 초성자 17, 중성자 11, 종성자 17 = 45자
- 45+4(다른 ㅈ)+2(Fc, Fj) 자 < 94 자
- 1 바이트 부호계

F ㄱ ㄴ ㄷ ㄹ ㅁ ㅂ ㅅ ㅇ ㅈ ㅊ ㅋ ㅌ ㅍ ㅎ
F ㅏ ㅑ ㅓ ㅕ ㅗ ㅛ ㅜ ㅠ ㅡ ㅣ
F ㄱ ㄴ ㄷ ㄹ ㅁ ㅂ ㅅ ㅇ ㅈ ㅊ ㅋ ㅌ ㅍ ㅎ

## 적합성 평가 1

### □ 코드와 문자표현의 범위

(코드)	(글자수)	(자소)
자소형(정음형)	약 399억	O
완성음절형	2,350	X
ISO 완성음절형	11,172	X
ISO 자소형	493,020	O(-)
자모형	11,172	O(-)
조합음절형	11,172	O(-)

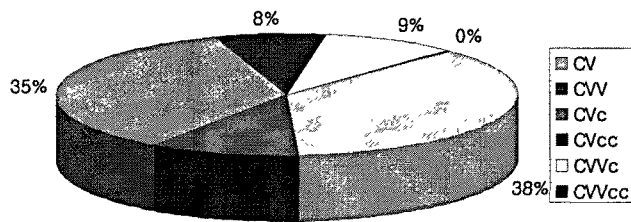
단기 4338년 6월 13일

동국대-변정용

31

## 적합성 평가 2

### 공간효율



C:초성, V:중성, c:종성

단기 4338년 6월 13일

동국대-변정용

32



## 구현

- 한글 과학과 공학
  - 컴퓨터에 가장 적합한 문자
    - 소프트웨어 <=> 글자 생성 원리
  - 정보화 시대에 최적의 문자
    - 일본 가나의 한계
    - 중국 한자의 고민
    - 로마자 타자에서 교타의 한계
      - QWERTY, Dvorak keyboard
  - 지적노동생산성 향상 분석
    - 한글
    - 한자, 로마자, 가나

단기 4338년 6월 13일

동국대-변정용

33

## 결론 -1

- 과학적인 문자 한글과 과학적 산물인 컴퓨터의 만남
  - 한글은 컴퓨터에 맞지 않는 문자라고 했다.
  - 훈민정음 해례는 과학적임을 입증하고 있다.
- 세계가 인정한 한글
  - UNESCO 세종상
    - 문맹퇴치 공로자에게
  - 세계의 문자로
    - 맥콜리 교수(미국) : 한글날 기념식
    - 중국에 권함 ... 한말 개일 목사
    - 볼리비아에서 실험 ... 북한

단기 4338년 6월 13일

동국대-변정용

34

## 결론-2

- 훈민정음의 과학성과 컴퓨터 과학을 일치시켜야 한다.
  - 유니코드 및 ISO 10646에 반영된 3 가지 한글 부호계의 공학적 타당성 검증필요
  - 과학적 검증 결과에 따라서 훈민정음 원리에 기반하고 있는 정음형을 기본 부호계로 해야 함
- 과학성에 입각하여 호환자모형 및 완성음절형 부호는 취소함으로써 국제 문자 공간을 절약하여 한자에 양도해야 한다.
- 한글부호 제정 및 표준화의 시행착오의 고리 끊어야 한다.