

고전 자료 디지털화사업에서의 신출한자 처리 개선방안

이 정 화
한국한의학연구원

Abstract

Improvement plan for 'Newly found ideographs(新出漢字)' in the digitalizing business of the old Korean documents

Lee Jeonghwa
Korea Institute of Oriental Medicine

As entering the information age of the 21st century, Korea is actively processing many digitalizing businesses related to information source of the Korean academic science at the government level based on the Korean advanced digital technologies, which makes them more evolved through the internet networks in Korea.

The definition of 'Newly found ideographs(新出漢字)' are made by researching and extracting from the old Chinese documents through the digitalizing process and they are not registered yet among the block of Unicode & extended Chinese characters those are existent international standard.

Presently Korea is in the middle of brisk developing computerized old documents in the huge scale. Meanwhile, the international standard of Chinese characters in mostly Asian countries where using them is processing and developing by IRG.

Therefore, Korean processing works is very important which are included extracting precisely 'Newly found ideographs' founded from building its database, organizing as an international standard code, submitting the International organization and finally registering as the best standard code.

Key words : Newly found ideographs, digitalizing old Korean documents, BMP, Extension A, Extension B, Extension C1, Extension C2

I. 들어가는 말

인류의 역사는 기록의 역사이다. 어떤 형태로든 과거의 지식과 사건들은 기록이 이루어졌으며 이것들은 후세에 전달이 되고 그것을 바탕으로 문명은 계승 발전된다. 우리민족은 15세기 훈민정음을 창제하여 오늘날 한글을 국어로 사용하고 있으나, 한글 창제 전과 그 후로도 얼마간은 한자가 공식적인 표기 수단이었다. 때문에 우리 조상들의 전통문화유산 중 80%이상이 한자로 기록되어 있으며 한글이나 기타 문자로 되어있는 기록은 아주 소수에 불과하다.

한자는 중국의 문자임에도 불구하고 오랜 기간 우리 민족에게 영향을 미쳤으며 한·중·일 삼국에서 공통 문자의 역할을 수행해 왔다. 그러나 한·중·일은 각기 독자적인 문자 정책을 이끌어옴으로써 그 쓰임과 자체가 조금씩 달라지게 되었고, 때문에 같은 형태를 띠고 있는 문자임에도 불구하고 다르게 쓰여 왔다.

오늘날 정보통신 기술의 발전은 문화유산을 전승하는 데에도 많은 변화를 가져오게 하였으며, 이것은 소실되어가고 사장되어가는 문화유산을 디지털화하여 데이터베이스로 구축하고 이를 웹상에서 서비스 하는 데까지 이르게 하였다. 특히, 이러한 정보통신의 발달에 힘입어 같은 한자문화권인 한·중·일 삼국은 역사와 문화를 활발하게 교류하기 위해 많은 노력을 기울이고 있다.

그러나 서구 주도하에 시작된 정보통신의 발달은 한·중·일 삼국에서 사용하고 있는 한자를 전산화하기에 그리 용이하지 않으며, 완벽한 구현을 하기

에는 많은 어려움이 있다. 컴퓨터는 기호에 의해 작동의 명령어와 각종 프로그램으로 실행을 하게 되는데 이러한 컴퓨터 언어가 영어로 되어있고 그 프로그램 체계 또한 영어에 맞게 개발되었다.

때문에 서구 중심의 컴퓨터 기술은 한자의 입출력을 배려하지 못하였으며, 이로 인한 디지털화의 한계는 여러 문제점을 나타내고 있다. 우리나라 경우도 마찬가지여서 지식정보자원관리사업이나 기타 다른 전산화 사업에서, 한자 또는 고어가 기록되어 있는 고전 자료를 디지털화하는데 있어 한자를 완벽하게 처리하기란 쉽지 않은게 현실이다.

이러한 상황을 개선하기 위해 국제적으로는 한자 문화권의 사용 한자 코드가 통합 정립된 국제 표준 문자 코드(Unicode¹⁾)를 완성하기 위해 노력하고 있다. 이 코드 체계가 완성되면, 국내외의 모든 문서 편집기는 이 코드 체계를 이용하게 될 것이며, 한자에 대한 정보를 처리하게 된다. 이를 우리는 국제 표준 문자 코드(UCS 10646-1)라고 부르는데, 이 코드를 부여받은 한자는 27,484자에 달하여, 한·중·일 각국의 한자는 각국의 사용 관례를 존중하여 그 고유 자체가 각각 달리 등재되어 있다.

본 연구는 신출한자²⁾의 등록과 처리절차에 대하여 살펴보고, 우리나라에서 현재 진행 중인 지식정보자원관리사업 중 한국학 고전 자료를 데이터베이스로 구축하는 사업인 역사정보통합시스템과 한의학 지식정보자원 관리사업 두 가지를 예로 들어 한국학 디지털화에 있어서의 신출한자 수집 현황과 사례를 살펴보고, 그 개선방안을 찾아보고자 한다.

1) 컴퓨터에서 세계 각국의 언어를 통일된 방법으로 표현할 수 있게 제안된 국제적인 문자 코드 규약. 국제표준으로 제정된 2 바이트계의 만국 공통의 국제 문자부호 체계(UCS : Universal Code System)를 말한다. 애플컴퓨터·IBM·마이크로소프트 등이 컨소시엄으로 설립한 유니코드(Unicode)가 1990년에 첫 버전을 발표하였고, ISO/IEC JTC1에서 1995년 9월 국제표준으로 제정하였다. 공식명칭은 ISO/IEC 10646-1(Universal Multiple-Octet Coded Character Set)이다.

2) 신출한자는 현재 확정된 표준 용어는 아니다. 사용하는 사람에 따라 신출자, 벽자, 비표준한자 등 여러 가지로 쓰이고 있다. 본고에서는 한문으로 이루어진 고전자료의 전산화 과정에서 새롭게 검출된 한자 중 전산처리가 어렵거나 불가능한 한자들의 의미로 '신출한자'를 사용하고자 한다.

II. 디지털화와 신출한자

21세기 정보화 시대로 접어들면서 우리나라에서는 앞선 디지털기술의 바탕위에 한국학 지식 정보 자원을 국가적인 차원에서 디지털화하는 사업들을 활발하게 진행하고 있다. 이는 발달된 인터넷 전산망을 통하여 더욱 더 활기를 띠게 되었으며, 축적된 데이터들은 전 세계에 서비스되고 있다.

우리나라의 고문헌 또는 기록물들은 대개 한자로 기록되어 있다. 우리들은 이를 잘 보존하고 올바르게 연구해야 할 의무가 있으며, 정보통신의 발달과 연계하여 한국학 고문헌을 디지털화하여 보존하거나 활용하려는 작업을 진행함에 있어서 원문을 그대로 이미지로 재현하거나 텍스트로 정리할 때 작업의 미숙이나 오류로 인한 원문의 훼손이 일어나지 않도록 많은 주의를 기울여야 한다.

그러나 고전자료의 디지털화뿐만 아니라, 한·중·일에서 사용되던 한문 자료의 통용이나 완벽한 재현은 그리 쉽지 않은 것이 현실이다. 특히 같은 한자이기는 하나 각 나라별로 문자 체계나 자체 또는 쓰임이 달라져 이를 공유하기란 현실적으로 쉽지 않다. 때문에 삼국에서는 이러한 문제점을 극복하고 각국의 국익과 문자의 편리성을 추구하고서 적극적인 대응을 하고 있는 것이다.

현재 우리나라는 대규모 고전 전산화 작업이 한창 진행 중이다. 그러므로, 데이터베이스 구축과정에서 발견되는 신출한자를 정확하게 검출하고 이를 ISO IRG(Ideographic Rapporteur Group, 표의문자코드전문가회의) 국제문자규격으로 정리하여 그것을 즉각적으로 국제표준기구에 제출함으로써 표준코드로 등록하는 작업은 매우 중요하다.

중국이나 일본의 경우는 이에 대한 적극적이고 활발한 연구가 이루어지고 있으며, 심하게는 미래에 쓰일 수 있는 문자에 대해서도 이에 대한 체계와 자체를 정리하여 유니코드로 등록하기 위해 노력을 하고 있다. 이에 반해 우리나라의 경우는 고전문헌의 재현뿐만 아니라 일상 문자생활에도 좋지 않은 영향을 미칠 수 있는데도 불구하고 지원과 연구가 많지 않으며 적극적인 대책 또한 세우지 못하고 있는 실정이다.

1. 국제표준규격과 한자코드

국제표준코드의 필요성은 정보통신기술의 발달로 다국어간의 정보 소통이 중요한 문제로 부각되면서 대두하였다. 국제 표준코드로서 유니코드(Unicode)가 실용화되면서 세계는 바야흐로 단일코드 시대에 접어들고 있다. 유니코드는 1996년 V2.0이 발표되어 일부 운영체제의 표준코드로 채택된 바 있고, 1999년에는 V3.0이, 2001년 5월에는 V3.1이, 2003년 DPSMS 4.0³⁾이 발표되었다. 향후 대부분의 운영 체제나 응용소프트웨어가 유니코드를 표준코드로 채택할 것이 확실시 된다.

특히 한자문화권의 경우 공통의 문화적 기반으로 인하여 각국 사이의 원활한 정보 소통의 필요성은 더욱더 절실하다. 이러한 필요성에 부응하여 현재 한자문화권 국가에서 사용되는 한자코드에 대한 국제 표준규격은 'ISO/SC2/WG2' 산하에 소속된 IRG를 통하여 진행되고 있다.

한·중(대만포함)·일은 표의문자코드전문가회의(IRG)를 결성하여 회합과 연구를 거듭한 끝에 당초 각국이 수용을 희망한 약 5만 4천자의 한자를 20,902

3) 유니코드 V3.0 한·중·일 통합한자(Super CJK)는 27,484자이며, 1991년 이후 추진되어 온 다국어 문자 처리를 위한 유니코드 표준(Unicode Standard)이 2003년 4월 버전 4.0.0으로 업그레이드되었다. 이번 업그레이드는 V3.2에서 업그레이드된 것으로 중요한 변화가 있다. 우리나라와 동아시아 관련된 추가 사항으로는 ① CJK 이체 선택자 블록 추가 : 향후에 등재될 CJK 이체자들을 위한 이체 선택자(block of variation selectors)라는 영역이 추가되었다. ② 사전 용도의 추분절음 부호 문자(약물)들이 추가되었다. ③ 동아시아 전각·반각 문자의 글자 폭에 관한 규정이 유니코드 표준 부속 문서에 삽입되었다.

개의 한자군으로 통합하는 최종안(ISO 10646-1)을 확정하였다. 여기에 1997년 Extension A가 추가되어 총 27,484자가 국제표준문자코드로 제정되었고, 이는 유니코드 V3.0에 수용되었다. 이후 2000년 12월 서울에서 열린 IRG 회의에서 Extension B 42,711자가 표준코드로 확정되었다. 이와 같이 표준코드로 제정된 CJK⁴⁾ 한자 가운데 현재 운영체제나 각종 응용프로그램 개발에서 실용화 된 것은 Unicode CJK Unified Ideographs Extension A 27,484자이다.

그런데 유니코드 한자는 한·중·일 통합 한자(Super CJK)를 기반으로 하였기 때문에 우리의 문자 관습과 차이가 날 수밖에 없다. 따라서 이러한 문제를 극복하고 유니코드 한자를 우리의 언어·문자생활과 정보 환경에서 보다 효율적으로 이용하기 위해서는 다양한 기초 연구가 이루어져야 할 것이다.

한 나라의 문자 정책은 쉽게 바뀔 수 있는 것이 아니다. 게다가 컴퓨터가 보급되면서 한·중·일의 서로 다른 전산 한자 자체와 부호체계로 인해 국제 인터넷 전산망을 통한 신속한 정보·문화 교류에 상당한 장애가 발생하고 있다. 현실을 고려할 때 한·중·일이 협의를 통해 한자를 표준화하는 작업을 하여야 하나 그리 간단치 않다. 이를 위해서는 먼저 우리 한자의 자체, 독음, 교육용 한자를 표준화함은 물론이거니와 동시에 컴퓨터에서 사용하는 한자도 표준화해야 한다.

2. 신출한자의 개념과 유형

신출한자는 한문 고전 전산화 작업 과정에서 새롭게 조사·추출된 한자로서, 기존의 국제표준규격인 유니코드 확장한자 문자세트에 아직 등록되어

있지 않은 한자를 말한다. 현재 한자문화권 국가에서 사용되는 한자 코드에 대한 국제 표준규격은 표의 문자코드전문가회의(IRG)를 통하여 진행되고 있다.

국제 문자 부호계에 포함된 ISO 국제표준한자는 다음과 같다.

- 기본 다국어 평면
(BMP, Basic Multilingual Plane) : 20,902자
- 한중일 통합한자 확장 A
(CJK Unified Ideographs Extension A) :
6,582자
- 한중일 통합한자 확장 B
(CJK Unified Ideographs Extension B) :
42,711자
- 합계 : 70,195자

따라서 신출한자의 범위는 위의 목록에 들어있지 않은 한자. 곧, Extension B 영역 이외의 한자를 말한다. 단, 우리나라의 경우 중국이나 대만과는 달리 범용 OS와 주요 워드프로세서에서 Extension A 영역의 한자까지만 입출력을 지원하기 때문에 Extension A 영역외의 한자부터 신출한자로 처리하고 있다. 다시 말해, Extension B 영역의 한자부터 신출한자로 처리하고 있는 것이다. 물론 소프트웨어를 사용하여 임시적으로 Extension B 영역의 한자를 입출력할 수는 있지만, 현재로서는 그것도 일부의 프로그램에서만 가능하다.

현재 ISO가 제정한 국제표준한자는 모두 70,195자이며, 지식정보자원관리사업을 통해 2만 여자⁵⁾ 정도의 신출한자가 발견 및 수집되었다.

신출한자를 유형별로 정리하면 다음과 같다.

4) CJK란 China(중국), Japan(일본), Korea(한국)를 가리킨다.

5) '99 ~ 2003년까지의 지식정보자원관리사업을 통해 발견된 신출한자는 2만 여자 정도이다. 이는 각 사업별로 자체 취합하여 판정한 것이므로 정확성은 떨어질 수 있다.

- * 이체자류
- BMP 이체자
- Extension A 이체자
- Extension B 이체자
- * 신출자류
- Extension B
- Extension C1 제안 한자(한국)
- 신출한자

위에서 보듯 신출한자는 크게 이체자(6)류와 신출자류로 구분할 수 있다. 이체자류는 기존 유니코드에 등록된 한자와 자형만이 다른 이체자들을 가리키고, 신출자류는 기존에 등록되지 않은 한자로서 새롭게 발견된 한자를 말한다. 이처럼 다양한 유형의 신출한자를 정확히 선별해내는 작업은 그리 간단치 않으며 전문가의 의견이 필요하다.

3. 신출한자 처리 절차

신출한자를 정확하게 판정하기 위해서는 전문가 관에서 여러 단계의 검정과정을 거쳐야 한다. 이는 무분별하게 수집하여 많은 양을 확보하기 보다는 가능한 정확한 목록을 작성하여 국제 표준기구에 제안하는 것이 바람직하기 때문이다. 국제적으로 공감하고 인정할 만한 신출한자를 제안하지 못하면 국제 표준으로 수용되기 어려울 뿐만 아니라, 이후 한자 표준화사업에 대한 국가적 신뢰를 잃을 수도 있으며 제안 자체가 거부되기도 하기 때문이다. 그만큼 신출한자 처리는 신중하고 정확하게 이루어져야 한다.

우리나라의 대표적 신출자 연구센터로 ‘문자코드

연구센터’⁷⁾가 있는데 이곳에서 이루어지고 있는 신출한자의 처리 과정을 보면 아래와 같다.

1) 원전 대조를 통한 오류 여부 검정

이 검정 작업은 문자코드연구센터에 제출된 신출한자 목록의 한자들을 원전과 직접 대조하면서 오독으로 인한 오류나 폰트 오류 여부 등을 조사하는 과정을 말한다.

(1) 원전 자형의 오독

원전 자형의 오독은 원전 한자의 자형에 대한 착각이나 이해의 부족으로 기존 표준코드에 등록되어 있는 한자인데도 불구하고 변별하지 못해 발생하는 경우이다. 이 같은 예는 한문 고전을 전산 입력할 때 사용되는 여러 가지 자동화 시스템들 가운데 세그먼트의 과정이라 해서 원전 데이터를 스캔 하듯이 이미지로 읽어 들여 글자 한 글자 한 글자를 박스로 처리한 다음, 글자를 개별적으로 인식하는 과정에서 나타나는 세그 즉, 박스처리의 오류에서 나타나는 현상이다. 위 또는 아래 글자의 획을 잘못 인식한다거나, 혹은 글자의 획이나 부분을 잘라내고 인식하는 경우이다. 이 같은 현상이 자주 나타나기는 하지만 대개는 교정 단계에서 수정을 하게 된다. 그리고 입력자의 전문성 또는 한자의 숙지도에 따라 그 발생 빈도에 가감이 있을 수 있다. 이 같은 오류가 국제 표준기구에 의해 발견될 경우는 국가

6) 이체자는 문자를 사용하는 과정에서 여러 가지 요인으로 발생한 正字와 音義는 같지만 字形이 다른 자를 가리킨다. 이체자는 혹은, 음의가 같지만 서로 字體가 다른 개념만으로 이해되어, 단순히 동일 음의를 가진 자체가 서로 다른 두 가지 관계만을 규정하는 용어로 해석될 수 있다. 그러나 일반적으로 이체자는 “某字의 이체자”라는 뜻으로 해석되어, 두 자체의 관계를 나타내는 것에 그치지 않고 정자에 대응하는 자체임을 나타낸다. 이체자라는 용어는 이와같이 동일자의 다양한 자체관계를 나타낼 뿐만 아니라, 바로 어떤 한자의 정자의 다른 字體라는 뜻을 함유하고 있다. 別字, 別體字, 異文, 字體之異, 僞體, 譌字, 謬體, 繆體, 俗字, 俗體 등으로 불리기도 한다. (『漢字 異體字 辭典 編纂 방법론 연구』『국제 문자 코드 제안 한자의 표준화에 대한 연구』, 李浚碩, 문화관광부, 1998.)

7) 문자코드연구센터는 1998년에 비표준문자 연구와 등록 지원을 담당하는 국가 공인 대표기관으로 출범하여, 비표준문자의 국제표준화 작업 및 문자코드 관련 제반 사업을 지속적으로 추진해 왔으며 2003년 비표준문자등록센터에서 문자코드연구센터로 명칭을 변경하여 오늘에 이른다.

신뢰도에 많은 타격을 입게 되므로 반드시 철저하고 정확한 점검이 필요하다.

(2) 폰트 자형의 오류

폰트 자형의 오류는 신출한자를 시스템상에서 구현하기 위하여 새로 각각의 글자를 새로운 폰트로 제작하는 과정에서 발생하는 문제로 획의 일부가 잘못 구현되는 경우에 많이 발생한다.

(3) 이체자 처리의 오류

이체자 처리의 오류는 이체자가 아닌데 이체자로 잘못 파악한 경우이다. 이는 원전의 한자가 오래된 목판이나 활자인 경우 자획의 일부가 떨어지거나 뭉개져 원 글자와 다른 글자로 인식하는 경우에 많이 발생하며, 또 다른 경우로는 관례적으로 또는 통상적으로 획을 간략하게 표현하는 경우 이를 다른 자로 인식함으로써 생기는 문제이다.

(4) 한자 정보의 오류

한자 정보의 오류는 해당 한자의 부수, 정보, 총획수 정보, 잔여획수 등을 잘못 파악하여 나타난다. 예를 들면 示(示)를 衣(衣)로, 入을 人으로 착각하는 등의 오류이다.

이같이 신출한자의 오류는 대개 원전 자료의 입력과 교정 등과 같은 1차 작업에서 주로 발생한다.

따라서 초기 작업 인력의 질적 수준에 의해 오류의 비율에 많은 차이가 나게 된다. 때문에 신출한자에 있어서 초기 작업은 매우 중요하다 하겠다.

2) 신출한자의 판정

(1) 신출한자의 처리 절차는 원전 대조를 통하여 오류 여부를 검정함으로써 문제가 없는 신출한자 리스트만을, 『Super CJK(Ver 14.0)』⁸⁾ 코드북과 대조하여 신출한자 여부를 판별한다. 이는 해당 한자의 부수와 잔여획수를 확인하여 수록 여부를 통해 신출한자 여부를 판정하는 것으로서 획수 순서의 배열이 모두 일치하지는 않으므로 $\pm 1\sim 2$ 획까지를 대조하여 정확하게 점검하도록 하여야 한다.

(2) IRG의 ‘Annex S(IRGN951)’⁹⁾ 문서 기준에 의거하여 이체자를 통합 및 선별하는 단계이다. ‘Annex S’는 유사 자형 한자와의 구분과 통합에 대한 기준이다. 이는 각국에서 이체자를 무분별하게 제안하는 것을 통제하기 위한 것으로서 기존 이체자를 기준으로 하고 이후 제안되는 신출한자에만 적용되는 기준이다.

(3) Ext. C1¹⁰⁾ 제출 한자 목록과 대조하여 Ext. C2 목록에 추가할 것인가 여부를 결정한다. Ext.

8) Super CJK Ver.14.0 검색 시스템은 문자코드연구센터 웹사이트(<http://ikc.korea.ac.kr/~cncsc/>)를 통해 기존에 구축하여 운영하고 있던 ‘한자정보 데이터베이스’와 ‘옛 한글 정보 데이터베이스’도 함께 서비스된다.

9) ‘Annex S(IRGN951)’문서는 IRG의 기술문서로서, 유사 자형 한자의 구분과 통합에 대한 기준이 제시된 문건이다. 이 문건의 취지는 기본적으로 각국에서 이체자를 무분별하게 제안하는 것을 통제하기 위한 것이다. 이 문서는 기존에 수용된 이체자는 그대로 두고 이후 제안되는 신출한자에만 적용되는 기준인데, 통합해도 무방하다고 판단되는 이체자와 구분해야 할 이체자를 예시를 통하여 기준을 정하여 이를 표준화하였다.

10) Ext. C1 제출한자 목록은 2002년 4월에 IRG에 제출한 우리나라의 신출한자 목록이다. Ext. C 영역의 한자는 C1과 C2로 나누어 제안하게 되는데, 현재는 각국에서 C1영역의 제출 한자 목록을 검토하는 단계이다. 따라서 향후에 추가되는 신출한자에서 Ext. C1 제출 한자 목록과 대조하여 중복되지 않는 한자를 선별하여 C2 영역의 제안한자 목록에 추가로 등록하는 것이다. 현재 IRG에서는 통합한자 확장 B영역까지 약 7만 여자의 한자를 표준코드로 제정하였으며, C1 작업에 대한 검토는 올해 11월 한국에서 개최될 IRG 회의에서 최종 확정될 예정으로 있다. 내년도에는 통합한자 확장 C2 영역의 등록 및 심의가 본격적으로 이루어질 것으로 보이며 이에 대한 준비가 시급한 실정이다.

C1 제출 한자 목록과 대조하여 중복되지 않은 한자를 Ext. C2 제안한자 목록에 추가로 등록한다.

(4) 신출한자로 검출된 한자에 대한 문자정보를 확인한다. 부수, 잔여획수, 총획수 등의 정보를 중심으로 재검토하는 단계로서 국제 표준 제안한자 양식에 맞게 필요한 정보를 확인하여 입력하는 과정이다. 이를 통해 신출한자 여부의 최종 판정 및 신출한자 목록을 작성한다.

신출한자의 표현 방안은 부호를 사용하는 방법, 자형의 조합식을 사용하는 방법, 한글음을 사용하는 방법 등이 입력 교정 단계에서 사용되고, 웹 화면 출력은 태그(Tag)를 사용하는 방법으로 이루어진다.

3) 신출한자의 국제 표준 처리 규정¹¹⁾

(1) 원칙

Ext. C1은 Super CJK(ver 14.0)와 통합할 수 있는 것이 분명한 한자로 한정하며, Ext. C2는 Super CJK(ver 14.0)와 통합할 수 있거나 'Annex T¹²⁾(N601 문서: IRG 近似字 통합 규정)'의 적용이 가능한 한자로 한정한다.

(2) 세부기준

모든 제출 한자는 'ISO/SC2/WG2(N875)'의 개정된 summary form에 의거해야 하며, IRG의 Unification rule(Annex T)을 Ext. C1에서는 엄격히 준수하여야 한다. 이 규정에 의해 의심되는 글자는

Ext. C2에 따로 제출해야 한다. 글자들은 Super CJK Ver 14.0과 충분히 대조하여 중복되지 않도록 주의하여야 하며 모든 글자는 128×128 Bitmap image(명조체)로 제출해야 한다. 모든 글자는 Kangxi index, 부수, 부수 제외 부분의 첫 획, 부수 제외 부분의 획수를 제시한다.

(3) 접수 거부 기준

제출 마감 시한 경과한 경우는 이를 거부하게 되며, Annex T rule의 적용의 오류가 5% 이상일 때 또는 IRG¹³⁾(N881)에 비추어 보아 부정확한 data format 일 때가 이에 해당한다.

Ⅲ. 역사정보통합시스템¹⁴⁾에서의 신출한자

1. 역사정보통합시스템에서의 신출한자

역사정보통합시스템 구축사업은 역사관련 4개 기관이 함께 참여하여 2004년 현재 5차 사업이 진행되고 있다. 본고에서는 신출한자 수집과 보고가 끝난 2003년 4차 사업까지를 대상으로 한정하여 살펴보기로 한다.

1999년부터 2003년까지 4차에 걸쳐 실시된 역사정보통합시스템에서 수집된 신출한자의 수는 총 4,419자이며, 신출한자가 발견된 문헌은 약 2,500책이다. 대략 한 책 당 2자 정도의 신출한자가 출현한 셈이다. 또 4차 사업까지 입력한 전체 책이 약 20,000

11) 정우봉, 박종우. '신출한자의 국제표준 처리 규정과 사례'. (2004년 문자코드연구센터 워크숍 기초발표문 【신출한자의 처리와 국제 표준화】. 문자코드연구센터. 2004.7.22.)

12) 'Annex T(N601)'는 IRG의 기술문서. 유사 자형 한자와의 구분과 통합에 대한 기준이 제시된 문건으로 이를 수정보완한 문서가 'Annex S(IRGN951)'이다.

13) 'IRG N881'은 IRG의 기술문서로서 새롭게 검출된 한자를 IRG에 제안하기 위한 제안한자 기술 포맷을 말한다. 즉, 이 기술문서 양식에 맞게 필요한 정보를 정확하게 기재하여 IRG에 제출함으로써 신출한자를 제안하게 되며 이 문서를 기준으로 심의를 거쳐 통합 한자를 확정하게 된다.

14) 역사정보통합시스템 구축사업은 역사 관련 기관 4개 기관(국사편찬위원회, 민족문화추진회, 한국정신문화연구원, 서울대 규장각)이 역사관련 지식정보자원을 통합적으로 서비스하기 위하여 구축하는 시스템이다.

책 정도이므로 총 대상자료 대비 신출한자 출현서적의 비율은 약 10% 정도이다.

역사정보통합시스템에서 수집된 신출한자 중 300여자가 Ext. C1에 등록될 예정이며, 나머지 신출한

자도 정리하여 Ext. C2에 등재하기 위해 노력하고 있다.

역사정보통합시스템에서의 신출한자 수집 방법과 처리 및 관리방법을 연차별로 살펴보면 아래와 같다.

사업 차수	신출한자 수집 방법	신출한자 관리 방법	비 고
2000년 1차 사업	· 상용 어플리케이션이 지원하지 못하는 한자를 웹상에서 구현하는 것에 주력	· Ext. B 코드를 구별없이 수집: Ext. A 한자 이외의 한자에 대한 웹서비스 문제에만 주력	· 신출한자 목록확보에만 주력
2001년 2차 사업	· 유니코드에 없는 한자의 수집과 이의 재활용에 초점	· Ext. B 한자코드에 대한 어플리케이션 지원에 대비하여 신출한자와 Ext. B를 용이하게 구분할 수 있는 인코딩 시키마 규칙 적용	· 역사정보통합시스템의 “신출한자 관리시스템”을 개발
2002년 3차 사업	· 유니코드에 초점	· 성균관대 동아시아학술원에 의뢰 - 신출한자 검정실시	· 1, 2, 3차 사업의 신출한자 모두 검정
2003년 4차 사업	· Ext. A와 Ext. B, C1을 구별하여 신출한자 수집	· 성균관대 동아시아학술원에 의뢰 - 신출한자 검정실시	· 4차 사업의 신출한자 검정

역사정보통합시스템에서 산출된 신출한자 처리 결과를 연차별로 살펴보면 아래와 같다.

2. 역사정보통합시스템에서의 신출한자 처리

1) 신출한자 선정부터 판정까지

DB구축단계에서 신출한자로 판단되는 글자 선정
 → 임시기호 부여와 함께 사업지측에서 1차 검증
 → 사업 기간 중 발생한 신출한자 후보 취합 → 전문기관에 신출한자 판정의뢰 → 최종 판정된 신출한자에 대해 코드 부여(역사정보통합시스템 사업에서 정한 9자리 표기 양식) → 임시기호를 신출한자 코드로 변환 → 관문에서 서비스되고 있는 신출한자 DB 업데이트(유니코드 확정 이후 변환 작업 필요)

사업 차수	신출한자	Ext. B	모호	합계	비고
1차 사업	784	852	0	1636	
2차 사업	1159	560	12	1731	
3차 사업	1598	569	28	2195	
4차 사업	878	242	26	1146	
합 계	4419	2223	66	6708	

2) 신출한자 코드 및 웹 서비스

현재 9자리 코드를 사용하고 있다. 앞의 두 자리는 부수, 뒤의 두 자리는 획수, 다음은 Ext. B에 속하는지를 보여주는 한자리 기호이다. 그 다음 네 자리는 Ext. B 코드값(Ext. B에 없으면 0000)으로 한다.

웹에서의 구현을 위해 실제 데이터에는 해당 신출한자의 이미지를 링크하고 있다. 즉, E109A0000이라는 코드의 신출한자를 표시하기 위하여 E109A0000.gif의 비트맵 이미지를 작성하고 E109A0000이라는 신출한자가 들어있는 데이터에 해당 이미지 링크를 걸어 웹 상에서 표현하는 방법이다.

IV. 한의학 지식정보자원 디지털화 사업¹⁵⁾에서의 신출한자

한의학 지식정보자원 디지털화사업은 2004년 현재 2차 사업이 진행 중이므로 본고에서는 2003년 1차 사업에서 산출된 신출한자에 한정하여 살펴보고

자 한다.

한의학지식정보자원 디지털화사업 1차년도에는 총 163종의 고문헌에 대하여 이미지 5,000면, 텍스트 87,000면을 디지털화하였다. 이것을 글자 수로 계산하면 15,417,181자의 한의학지식정보자원을 디지털화 하였는데 여기서 산출된 신출한자 수는 『의방유취』에서 448자, 『의방유취』를 제외한 나머지 162종의 고문헌에서 469자가 발견되었다. 이로써 한의학 지식정보자원관리 1차 사업에서 산출된 신출한자는 총 917자¹⁶⁾이다.

917자의 신출한자는 고문헌 163종에서 산출된 것으로서 이는 고문헌 1종당 5.6자의 신출한자가 발견되었음을 알 수 있으며, 이를 다시 87,000면의 고문헌으로 계산하여 보면 100면당 약 1자 정도의 신출한자가 출현하였음을 알 수 있다.

1. 신출한자 처리 과정과 대상자료별 신출한자

한의학 지식정보자원 디지털화사업에서의 신출한자 처리 과정을 자세히 살펴보면 아래의 도표와 같다.

15) 한의학 지식정보자원 디지털화사업은 한의학 지식정보자원 웹서비스(<http://jisik.kiom.re.kr>)를 통해 서비스되고 있다.

16) 이는 신출한자 검정기관으로부터 판정을 받는 것은 아니므로 다소의 가감이 있을 수 있다. 한의학 지식정보자원 디지털화 1차 사업에서의 신출한자 판정은 사업자가 주관기관에 신출한자 목록을 제출하여 주관기관에서 확인을 하였을 뿐이다, 2차 사업부터는 검정기관에 의뢰하여 신출한자에 대한 명확한 검증을 할 계획이다.

단계		처리내용	비고
입력	최초입력	<ul style="list-style-type: none"> · 입력 단계에서 입력 불가능한 자들을 추출 - 한자의 정확한 판독이 요구됨 - 이체자는 정자로 입력 	<ul style="list-style-type: none"> · 신출예정자 관리테이블 기록 (엑셀시트에 목록화)
	입력불능 한자출현	<ul style="list-style-type: none"> · 마크업 부가 혹은 신출자 관리 번호 부여 · 신출예정자 관리테이블 기록 	
교정	마크업 발견	<ul style="list-style-type: none"> · 원본한자의 정확한 교정 · 이체자는 정자로 교정 	<ul style="list-style-type: none"> · 교정단계에서 정확성을 높인다
수정 입력	마크업 발견	<ul style="list-style-type: none"> · 이체자는 정자로 수정입력 · 정자는 정확한 자형대로 입력 · 입력불가능한자 집계, 리스트업 	<ul style="list-style-type: none"> · 신출자관리테이블 갱신
신출자 1차 정의	리스트 검토	<ul style="list-style-type: none"> · 이체자 / 정자 판정 · 신출자 여부 임시 판정 · 신출자의 임시 부가정보 생성 · 신출자관리테이블 갱신 · 임시코드 배정 	<ul style="list-style-type: none"> · 신출자 관리시스템 적용 · 신출자 관리위원회 운영
신출자 판정	신출자 판정	<ul style="list-style-type: none"> · 신출자 목록에서 중복자 삭제 · 역통, 규장각, 비표준문자센터 등 기존 신출자들과의 대조를 통해 확인 · 이미 존재하는 문자는 기존 코드값 부여 	<ul style="list-style-type: none"> · 엑셀시트에 표시 · 주관기관 신출한자 판정
코드 부여	임시코드 부여	<ul style="list-style-type: none"> · 최종 신출자를 대상으로 임시코드값 부여 · 부수/ 획수 순으로 코드값 부여 · 기존 문자는 해당 이미지 폰트 활용 	<ul style="list-style-type: none"> · 엑셀시트로 작업
신출자 제작	리스트 적용	<ul style="list-style-type: none"> · 코드배정, 폰트제작 · 유니코드 테이블 및 부가정보 테이블 업그레이드 	<ul style="list-style-type: none"> · XML Editor상에서 최종 산출물인 XML 문서 업데이트 · 신출자관리테이블 갱신
폰트 제작	이미지 폰트 제작	<ul style="list-style-type: none"> · 임시코드값이 부여된 신출자를 대상으로 이미지 폰트 제작 	<ul style="list-style-type: none"> · 엑셀시트에 이미지 폰트명 표시 · 16×16, 32×32 두 가지 사이즈로 제작
	서비스 반영	<ul style="list-style-type: none"> · 이미지 폰트 제작이 완료된 신출자는 XML 데이터에 이미지 링크를 점검한 후 서비스에 반영 	
	트루타입 폰트 제작	<ul style="list-style-type: none"> · 최종 신출자를 취합하여 트루타입 폰트 제작, 주관기관에 제출 	<ul style="list-style-type: none"> · 서비스에는 반영하지 않음

한의학 지식정보자원 디지털화사업에서의 대상자료별 신출한자 현황을 살펴보면 아래의 도표와 같다.

대상자료명	신출한자수	대상자료명	신출한자수	대상자료명	신출한자수
家庭救急方	2	婦人大全良方 9	1	宜彙	58
簡易辟瘟方	1	四時纂要抄	2	餌方	4
檢要	5	四醫經驗方	4	人夫須知	1
經驗痘方	1	生産秘方	8	仁濟志 1	3
經驗方	7	壽民妙詮	1	仁濟志 2	2
谷靑冗語	1	壽世寶訣	3	仁濟志 3	9
廣濟秘笈	11	壽養叢書類輯	3	仁濟志 4	17
救急簡易方診解	2	袖珍經驗神方	3	仁濟直指方	9
救急單方	8	食物本草	5	濟生篇	3
救急方	2	身機踐驗 上	16	舟村新方	3
及幼方 1	2	身機踐驗 下	11	增補山林經濟	2
及幼方 2	1	神應經	1	增修無冤錄診解	2
丹谷經驗方抄	3	新註無冤錄	4	纂圖方論脈訣集成	1
單方秘要□經驗新編	3	新集御醫撮要方	1	察病用藥訣	1
東西醫學要義	3	新纂辟瘟方	3	瘡疹集	2
麻疹奇方	1	新編集成牛馬醫方	12	草堂遺訣	1
幕經	2	兩無神編 1	2	春鑑錄	7
脈訣理玄秘要	1	兩無神編 2	1	鍼灸極秘傳	1
名方類證醫書大全	5	良方金丹	3	鍼灸集成	5
明善錄	1	診解臘藥症治方	4	鍼灸擇日編集	2
辟瘟新方	2	診解痘瘡集要	1	胎產要錄	1
葆養志	5	川藥賦	3	太醫院先生案	3
本草附方便覽 1	12	醫科先生案	14	八陣方	1
本草附方便覽 2	4	醫科八世譜	2	漢方醫學小兒專科	4
本草附方便覽 3	14	醫方綱要	1	解惑辯疑	7
本草備要	3	醫方類聚	448	鄉藥救急方	2
本草類涵要領	2	醫方集略 1	2	鄉藥集成方 2	4
本草精華	9	醫方集略 6	1	鄉藥集成方 3	6
本草彙英 1	4	醫方撮要	1	鄉藥集成方 4	4
本草彙英 2	6	醫方合編 1	5	鄉藥集成方 5	31
本草彙英 3	1	醫方合編 2	4	鄉藥採取月令	1
本草彙英 6	1	醫方合編 3	1	活人心法	1
本草彙英 7	3	醫寶	9	回春方	3
婦人大全良方 3	2	醫本	3		
婦人大全良方 7	2	醫藥鑑	1		

3. 한의학 고전자료에서만 나타나는 신출한자 관련 특이사항

한의학 고문헌에서 자주 쓰이는 특별한 약자나 속자 또는 간자의 경우 그 이해도에 따라 판단의 오류가 있을 수 있는데, 한의학 문헌에서 나타나는 특수한 용례를 이해하지 못하여 다른 글자로 판단한다거나, 또는 신출한자로 규정하는 오류를 범해서는 안 될 것이다. 이는 고전자료를 디지털화하여 우리의 문화유산을 보존하고 이를 연구하고 활용한다는 기본 취지에 반하게 될 뿐만 아니라, 문화유산을 해치는 결과를 초래하게 된다.

이러한 예는 한의학 주제뿐만 아니라, 다른 주제나 전공의 경우에도 마찬가지로 발생할 수 있는 문제점으로 각 주제마다 이를 충분히 조사하고 연구하여 정리할 필요가 있다. 이러한 작업들은 한 전문분야에서 통용되는 주관적이고 관습적인 경험의 지식들을 객관화하고 과학화 할 수 있는 좋은 계기라 여겨지며, 이를 취합하여 표준화하고 정보화하여 공유함으로써 인접 학문이 서로 발전할 수 있을 것이며 새로운 부가가치를 창출할 수 있을 것이다.

한의학 문헌 자료에서만 나타나는 한자의 특수한 용례를 몇 가지 정리하면 아래와 같다.

- 干(방패 간), 召(부를 소) : 薑(생강 강), 棗(대추 조) 대신 사용
/ 예) 강삼조이(薑三棗二 → 干三召二)
- 姜(성 강) : 薑(생강 강) 대신 사용
/ 예) 생강(生薑 → 生姜)
- 叵(돌아갈 귀) : 歸(돌아갈 귀) 대신 사용
/ 예) 당귀(當歸 → 当叵)
- 卜(점 복) : 蔔(무우 복) 대신 사용
/ 예) 나복자(蘿蔔子 → 蘿卜子)

- 古(옛 고) : 膏(기름 고) 대신 사용
/ 예) 석고(石膏 → 石古)
- 角(뿔 각) : 殼(껍질 각) 대신 사용
/ 예) 지각(枳殼 → 枳角)
- 別(나눌 별) : 鱗(자라 별) 대신 사용
/ 예) 별갑(鱗甲 → 別甲)
- 痲(악창 응) : 癰(종기 응) 대신 사용
/ 예) 응저(癰疽 → 痲疽)
- 戈(창 과) : 錢(돈 전) 대신 사용
- 三(석 삼), 參(석 삼) : 蔘(인삼 삼) 대신 사용
- 干(방패 간) : 乾(하늘 건) 대신 사용

위와 같은 용례는 단순히 웹에서의 구현 문제뿐만 아니라, 검색의 문제, 오해와 오독의 문제 등 많은 문제점들을 지니고 있다. 이에 대한 체계적인 정리와 연구가 필요하다.

V. 결론

신출한자와 관련하여 고문헌 자료를 디지털화하는 지식정보자원관리사업 중 두 가지 사업을 사례로 들어 살펴보았다. 역사정보통합시스템, 한의학지식정보자원 디지털화 사업 모두 한자 자동 입력 시스템에 의해 원문이 디지털화되고 있는데, 역사정보통합시스템의 경우는 신출한자 판정을 검정기관에 의뢰하고 있고 한의학지식정보자원 디지털화 사업의 경우는 사업자와 주관기관에서 임의로 선정하고 있다. 이 두 가지 사업의 예 뿐만 아니라, 다른 사업도 이와 비슷한 상황으로 신출한자의 관리가 제각각 이루어지기 때문에 정보의 공유와 표준화 또는 객관화가 어려운 현실이다.

신출한자 수집과 처리 그 자체가 지식정보자원관리사업의 목표는 아니다. 그러나 이러한 지식정보자원을 디지털화하여 웹상에서 완벽하게 서비스하기 위한 중요한 과제이기도 한다.

그러므로 표준한자코드를 제정하는 국가기관과 상호협력체제 하에서 효율적인 신출한자의 수집과 관리가 이루어져야 할 것이며, 어느 누구의 이익도 아닌 우리 모두의 이익 즉, 국익을 위한 움직임으로 이어져야 할 것이다.

한·중·일 삼국 중 중국과 일본은 이를 위한 많은 작업들을 하고 있으나, 우리나라는 이에 대한 연구와 적극적인 대응이 부족한 상태이다. 어느 한 기관이나 센터의 헌신적인 노력만을 기대할 것이 아니라 적절한 투자와 지원이 뒤따라야 한다. 원활한 정보이용을 위하여 일상생활에서 사용하는 문자와 한자가 유니코드에 없음으로 인하여 입출력이나 재현이 문제가 있게 해서는 안 되며 이를 주관하고 검정할 대표기관을 하루빨리 선정하여 국익에 우선하는 국가적인 차원에서의 대응과 적극적인 조치가 필요하다.

지식정보자원관리사업의 전담기관인 전산원에서는 정보화사업에서 산출된 신출한자에 대하여 조사하고 이를 검정하여 IRG에 제출하는 과제를 계획하고 있다. 그러나 이 계획은 1999년부터 2003년까지 산출된 2만 여자의 신출한자와 2004년 12월 사업이 종료되는 당해년도 사업까지를 포함하여 2-3개월 안에 이를 취합하고 검정하는 연구과제이다.

각 사업에서 문제점으로 지적되어 필요에 의해 진행되는 연구과제인데도 불구하고 아직은 미흡한 부분이 적지 않다. 먼저, 4~5년간에 걸쳐 산출된 2만 여자의 신출한자를 2달에 걸쳐 검정하고 판정한다는 것도 무리이고, 2004년 사업이 대개는 11월말 종료되는데, 이를 한달에 연구기간을 두고 12월 말까지 취합 정리하여 판정하다는 것은 시기적으로 무리한 작업일정이다. 게다가 어떤 정해진 원칙이나 기준 또는 이를 판정할 기관이나 센터가 정해지지 않았는데, 이를 정리 취합하여 판정하는 과제를 진행한다는 것은 문제가 있다. 연구 과제를 수행하

기에 앞서 좀더 거시적이고 장기적인 안목에서의 계획과 연구과제 사업의 진행을 기대한다. 이는 멀게는 국학 연구의 현대화를 통해 전통 문화를 재창조할 수 있는 실질적인 구현 기반으로 작용하게 될 것이며, 가깝게는 국어와 한문 교육 방법을 유효하게 개선하여 올바른 국어 생활을 영위할 수 있는 효과를 기대할 수 있을 것이다.

지식자원 관리사업에서의 신출한자 처리에 대한 개선 방안을 몇 가지로 요약 정리하면 아래와 같다.

1. 신출한자 처리 시기의 변경

지식정보자원관리 사업 작업 지침 상 신출한자는 사업기간 내에 처리하도록 되어 있으며, 또한 유관 사업의 결과물과 대조하여 신출한자 판정을 하도록 되어 있다. 그러나 사업진행상 모든 디지털화 공정이 끝나야 신출한자 처리가 가능하며, 더욱이 사업간의 신출한자 산출 시점이 각기 다르므로 이에 대한 변경이 필요하다. 바람직한 대안은 사업종료일까지 신출한자 목록을 사업자나 주관기관에서 지식정보자원관리 사업의 전담기관인 전산원에 제출하여, 종료 후 6개월 이상의 시간을 가지고 일괄로 처리하고 판정한 뒤 각 사업별로 통보하고 이를 서비스에 반영하도록 하는 것이 이상적이다.

2. 신출한자 판정 방법의 변경

현재는 사업자들이 임의로 신출한자 판정을 하게 되어 있고, 또한 신출한자의 기준도 각 사업마다 기관마다 조금씩 다르게 규정되어 있다. 따라서 우선 신출한자의 기준을 정확하게 정하여 정확한 신출한자 판정이 이루어지게 하여야 한다. 신출한자 판정은 KIRG¹⁷⁾나 비표준문자센터 또는 성균관대 동아

17) 한국한자특별전문위원회(KIRG: Korea Ideographic Rapporteur Group)

시아학술원 같이 충분한 인력과 경험을 갖춘 기관을 공식 검정센터로 선정하고 그곳에 신출한자 판정을 의뢰하여 검정을 받도록 하여야 한다.

3. 신출한자 코드 부여 절차의 변경

현행 엔코딩 스킴은 1차적으로 사업자가 각자 판단하여 부여하고 있고, 이 때문에 각 기관별로 신출한자 코드가 다른 결과를 낳고 있다. 이에 대해서 공식적인 기관에서 제시하는 일치된 엔코딩 스킴을 준수할 필요가 있으며, 이를 공식화하여 적용하는 것이 바람직하다. 가령 부수-획수-한글음가 순의 코드 부여를 통한 조합과 같이 전문가가 아니더라도 충분히 숙지할 수 있는 규칙과 내용에 대한 합의가 필요하다.

4. 신출한자 폰트 적용의 표준화

현재 지식정보자원관리 사업의 경우, 신출한자는 트루타입 폰트를 제작하여 주관 기관에 제출하는 경우가 대부분이다. 그러나 실제 웹 서비스에서는 이 트루타입 폰트를 이용하지 않고 이미지 폰트를 기반으로 서비스한다. 이처럼 웹 서비스를 위해 트루타입 폰트를 새롭게 제작하였음에도 불구하고 이용하지 못하는 이유는 이용자의 폰트 다운로드의

문제, 또는 윈도우 시스템 및 워드 프로세서와의 비호환성 등 여러 가지 문제가 있기 때문이다.

이에 대한 대안으로는 폰트제작에서 웹 서비스에 이르기까지의 과정을 표준화하여, 주관 기관에 제출하는 폰트 형태와 웹 서비스에 사용하는 폰트 형태를 일치시키는 시스템을 갖추는 방안을 마련해야 한다.

5. 신출한자 통합 관리

지식정보자원관리사업에서 뿐만 아니라, 한국학 자료를 디지털화하는 사업을 진행하는 과정에서 나타나는 신출한자에 관하여 이를 통합적으로 관리하고 표준화할 수 있는 시스템의 구축이 시급하다. 신출한자의 통합관리를 통한 취합 목록은 ISO IRG 국제문자규격으로 정리하여 국제표준한자로 등록할 수 있는 필수자료로 활용하게 된다. 이는 국제표준 규격으로 정리된 신출한자를 한국학 자료의 디지털화에 사용하게 됨으로써 DB의 품질을 높일 수 있을 것이며 각 사업별로 소용되고 있는 비용을 절감하는 경제적 효과도 거둘 수 있을 것이다.

검색어 : 신출한자, 고전자료, 디지털화사업, 지식정보자원, 한자코드

참 고 문 헌

1. 서경호. 『국제 문자 코드 제안 한자의 표준화에 대한 연구, 상·하권(연구보고서)』, 서울. 문화관광부, 1998.12.
2. 이재훈. 『국제 문자 코드계의 한자 표준화에 대한 연구(연구보고서)』, 서울. 문화관광부, 2001.12.
3. 정우봉. 『문자코드 표준화 연구 : 21세기 세종계획』, 서울. 문화관광부·국립국어연구원, 2003. 11.
4. 고려대학교 민족문화연구원 문자코드연구센터. 『신출한자 처리관련 업무 협의 - 회의자료』, 서울. 同센터, 2004.6.
5. 고려대학교 민족문화연구원 문자코드연구센터. 『신출한자의 처리와 국제 표준화 - 2004년 문자코드연구센터 워크숍』, 서울. 同센터, 2004.7.
6. 민족문화연구원 문자코드연구센터. 문자코드연구센터 소식 제13호, 고려대학교 민족문화연구원, 서울. 同센터, 2004.6.