

# 개방형 몽골어 웹 대사전 편찬 경과와 차후 과제

성비락(Sainbilegt Dashdorj)  
몽골국립대학교 아시아학과 교수

## 1. 들어가며

언어 사전은 무엇인가? 주지하다시피 ‘언어를 모아 일정한 순서로 배열하여 의미, 주석, 어원, 품사, 다른 말과의 관계 따위를 밝히고 풀이한 책’이라고 정의할 수 있는데, 크게는 백과사전과 어학 사전 그리고 전문용어사전으로 분류한다. 4차 산업 발전에 따라 빅데이터와 말뭉치, 자연 언어 처리와 자동화 등 다소 생소한 단어들이 등장하면서 예전에 사용했던 종이형 사전 형태가 하나의 큰 웹 기반 디지털 형태로 바뀌고 있으며, 사전 사용 방식에서도 다양한 모습을 보이고 있다. 즉, 사전 편찬 주체의 경계가 사라지는 현상을 크게 ‘자동화’와 ‘대중화’라는 두 가지 핵심어로 요약이 가능하다 (배연경, 2016:44)고 할 정도로 사전 편찬 모습이 달라지고 있다.

이와 같은 사전학의 세계적인 흐름을 살펴보자면 몽골어 사전 편찬 작업의 과거와 현재 그리고 미래에 대한 고민을 하지 않을 수 없다. 즉, 이 글에서는 몽골어 사전 편찬 중 우선 모국어 사전의 변천 과정을 살펴보고, 최근 많은 호평을 받고 있는 ‘개방형 몽골어 웹 대사전’의 편찬 경과와 차후 과제를 제시하려 한다.

## 2. 역사적 배경

몽골어 사전학 연구 성과에 따르면, 몽골어 사전 편찬에 대한 가장 오래된 기록은 서기 4세기경으로 거슬러 올라간다.<sup>1)</sup> 1921년까지 편찬된 사전들을 살펴보면 주로 국가적 차원에서 편찬한 사전과 종교적 목적으로 편찬한 사전으로 분류되는데 국가적 차원에서 편찬한 사전은 다시 모국어 사전과 다종의 외국어 번역 사전으로 나뉜다. 모국어 사전들은 백과사전과 모국어 대·소사전으로 나뉜다.

백과사전은 주로 갈래(분류) 사전 형태로 편찬되었는데 현대에 최초로 몽골인이 편찬한 사전은 ‘Guchin zurgat toil bichig(36권 사전)’이다. 이 사전은 2종으로, 하나는 푸르브(Purev)의 ‘44권 백과사전’이고 다른 하나는 바트오치르와 미식(Bat-Ochir, Mishig) 등의 몽골 학자들이 1921~1927년에 편찬한 ‘36권 백과사전’이다. 이 백과사전은 19,000여 개 단어를 수록·분류하여 뜻을 풀이하였는데, ‘몽골어와 만주어 사전과 기타 3개, 4개의 언어 번역 사전들의 기반이 되었다.

모국어 사전들은 표준어 규정을 지켜야 했기 때문에 방언형과 속어는 수록이 불가능했다. 한때, 활발하게 문자 생활을 하는 몽골 지식인들 사이에서는 ‘тольтой үсэг(국어사전에 수록된 단어라는 의미)’라는 이름이 붙은 단어들만 사용해야 사람들에게 인정을 받던 시기가 있었다(Б. Ринчен, 1964:124). 몽골에서 모국어 사전 편찬의 역사적 변천 과정을 좀 더 심도 있게 살펴봐야 하겠지만, 이 글에서는 개방형 몽골어 웹 대사전 편찬에 바탕이 된 현대 몽골어 사전의 역사적 배경만을 소개하고자 한다.

---

1) 기원 후 4세기 선비 시대에 출간된 13권 서책 목록이 나오는데 그중 여섯 번째 책이 중국어로 國語彙編物名三券 kouo-yu tsa wou-ming san kuan(국어 갈래 사전, 3권)이다. 이에 대한 소개가 L.Liget, «Le Tabghatch, un dialecte de la langue Sien-pi, Mongolian studies», Budapest, 1970. 19-20쪽에 나오는데 여기서는 Л. Шагдарсүрэн(2017: 152)에서 재인용한다.

첫째, 사전학 연구자들은 사도 바야스갈랑 삭즈(성: Sadoo 부친명: Bayasgalan 이름: Shagj)가 1929년에 편찬한 'Eriküy-e kilbar bolγan jokiyaγsan mongγol ügen-ü taylburi čindamuni-yin erike(쉽게 찾도록 편찬한 몽골어 소사전)'라는 사전을 현대 모국어 사전의 기저로 본다. 비록 이 사전의 표제어가 16,000여 개에 불과하고, 1994년에 이르러 비로소 출판이 되었으나<sup>2)</sup> 이 사전은 국내외 전문가들에게 널리 사용되었다. 이 사전은 그전에 분야별 갈래 사전들이 주를 이뤘던 시기의 사전 편찬 방법과는 달리 알파벳순으로 단어를 배열하여 독자들이 단어를 쉽게 찾을 수 있도록 했으며, 18~20세기에 편찬된 여러 사전들을 참고하여 집필했다는 점에서 몽골어 어휘론 연구와 표준 정서법 정착 및 발전에 크게 기여한 사전이라는 평가를 받고 있다.

둘째, 1966년에 편찬된 야담자브 체벨(부친명: Yadamjav 이름: Tsevel)의 'Mongγol helnii tovč tailbar toil(몽골어 풀이 소사전)'이 있다. 이 사전에 수록된 표제어는 30,000여 개 정도이며, 15년간(1949~1964년)의 노고 끝에 출판되었다. 이 사전의 서론에서 필자는 집필 장식과 의의에 대하여 다음과 같이 밝혔다.

- 1) 이 사전 편찬 목적은 현대 몽골어를 배우는 일에 도움이 되고자 한 것이다. 이에 몽골인들의 실생활어, 경제와 사회 정치 문화 관련어, 국가 발전에 의해 도입되는 시사어와 과학 용어들을 최대한 포함시키려 했으며 현대 문학과 역사책을 이해하는 데 필요한 옛날 단어와 구어를 두루 수집·수록했다.
- 2) 이 사전은 대사전이 아니므로 어휘 수보다는 동의어와 다의어 수록에 집중했으며 지금은 아예 안 쓰게 된 단어나 필요성이 적은 단어 및 지명 등은 제외시켰다.

2) 1994년에 몽골문화재단 지원을 받아 체벨소드놈(D. Tsevelsodnom), 멘드오요(G. Mend-Oyo) 등이 본 사전 원고를 북경에서 출판한 바 있다(Г. Мэнд-Оёо, 2013).

3) 이 사전 편찬 과정에 여러 사전들, 문학 작품들, 신문과 구어 자료들을 활용했다.

이 사전은 편찬 목적과 방식에서 기존의 사전들과는 구별되는 차별성을 갖고 있었을 뿐만 아니라 언어학 이론과 사전 편찬 방법론에서도 진보가 있었다(Л. Болд, 2008:15). 또한 2008년 몽골어 대사전이 나오기 전까지 대중적인 활용도가 높았다. 2013년에는 그의 아들인 샷드르수앵(Ts. Shagdarsuren) 박사가 이 사전을 증보하여 재출간했다.

셋째, 2008년에 집필이 완성된 ‘몽골어 대사전 5권은 개방형 몽골어 웹 대사전의 기저가 된 사전이다. 총 표제어가 80,000여 개이며, 12만여 개의 관련어와 용례를 수록한 대사전이다. 이 사전 집필은 1933년부터 시작되었다고 보는데, 그 이유는 몽골인민공화국(몽골국의 옛 이름) 총리의 1933년 1월 6일 제4호령에 의해 경서원(지금의 몽골과학아카데미)에서 연구원 7명으로 구성된 모국어 사전 편찬 위원회가 설립되었기 때문이다. 하지만 1933년부터 1937년까지 4년간 활발하게 이뤄졌던 사전 집필 작업이 당시 정치적 사건의 소용돌이 속에서 중단되었다. 앞서 소개한 체벨 선생은 대사전 집필이 지연되는 것과 모국어 사전 수요가 증가하는 것을 감안하여 소사전부터 만들어야겠다고 결심하였는데, 이것이 몽골어 소사전 편찬 동기이다(Я. Цэвэл, 1966).

이렇게 집필이 지연되었던 몽골어 대사전이 1980년부터 다시 빛을 보기 시작했다. 몽골 과학아카데미 선임 연구원 룽상발당(Kh. Luvsanbaldan)이 몽골어 대사전 편찬 방식을 새로 정립하고 몽골과학아카데미 어문연구소 연구원들을 동참시켜 대사전 편찬 작업을 재개했다. 그러나 1990년대에 몽골의 사회 체제가 바뀌면서 사회 전체가 혼란스러운 시기를 겪게 되면서 대사전 편찬 작업은 중단되었고, 이후 2005년에 이르러서야 다시 이어졌다.

몽골국 대통령 앵흐바야르(N. Enkhbayar)는 2005년 9월 26일 제46호령을 선포했는데 이는 ‘몽골어 대사전 편찬에 관한 것이었다. 1년 안(2006년)에 몽골어 대사전 편찬 작업을 마무리하도록 몽골과학아카데미와 그 산하 연구소 소장들, 그리고 관련 대학교 총장들을 대상으로 법령을 내린 것이다. 그 결과 몽골어 대사전 집필자들이 속도를 내서 다양한 프로젝트를 수행하여 대규모 자료 수집과 인적 자원 구축 및 관련 학술 활동을 전개한 후에야 2008년에 몽골어 대사전 5권이 완간되었다.

앞에서 말한 몽골어 대사전의 주요 특징을 정리하면 다음과 같다.

- 1) 근대 몽골어의 형성이 17세기 중순부터 시작되었다고 보는 시각에 따라 근대 이후 활용 빈도가 높은 어휘를 중심으로 자료 수집이 이루어졌다.
  - 오래 전부터 수집해 온 몽골 주요 문헌 및 문학 작품에서 뽑은 5백만 개의 단어 카드 활용
  - 다종의 몽골어 및 외국어 사전과 전문용어사전 활용
  - 1990년 이후 출간된 4개의 주요 일간지 활용
  - 그 외에 필요하다고 판단되는 문학 작품과 신문·잡지 및 구어 자료 활용
- 2) 몽골어가 합성어, 연어(連語), 관용어가 풍부함을 고려하여 관련한 용례를 제시하는 데에 주력했다.
- 3) 기존의 몽골어 사전들(B. Shagi, Ya. Tsevel, Ts. Norjin 등)에서 제시한 것과는 달리 동사의 피동형과 사동형, 동사의 상(aspect)을 나타내는 어미가 접속된 형태를 표제어로 수록했는데 단순한 문법적 설명만 하지 않고 좀 더 구체적인 뜻풀이를 했다.
- 4) 몽골어 표준 정서법 규정을 준수하도록 했다(Л. Болд, 2008:17-20).

### 3. 개방형 몽골어 웹 대사전 편찬 경과

‘개방형 웹 사전’은 그전에 편찬된 사전 중에 정보량이 가장 많고, 민간의 참여로 계속 변하는 말을 지속적으로 담아낼 수 있는 ‘개방형’이라는 틀을 가지고 있는 사전을 가리킨다(이현주, 2016:82-83).

2014년 12월 3일 몽골 대통령 비서실 부장인 네르구이(B. Nergui)와 몽골 과학아카데미 어문연구소 소장 볼드(L. Bold) 간에 업무 협약을 체결하였다. 이것은 몽골국 대통령 차히아 엘벡도르지(부친명: Tsakhia 이름: Elberdorj)의 구상(initiative)과 대통령 비서실의 지원으로 2014년 12월부터 2015년 12월까지 1년간 활발하게 이뤄진 ‘몽골어 웹 대사전’ 편찬 사업의 시작이 되었다.

이 사업의 책임자인 볼드는 몽골과학아카데미 어문연구소와 물리기술 연구소의 정보학과 연구원 11~13명으로 사업팀을 구성하고, 2014년 12월 17일에 개방형 웹 기반 몽골어 대사전 개발 착수 행사를 했다. 협약한 대로 1년 안에 마무리해야 하는 사업이었기 때문에 2015년 2월 12일에 사업 결과물을 공개했다. 60,000여 개의 표제어 용례, 그리고 80,000여 개의 관련어들(총 14만 개)이 입력된 개방형 웹 사전의 첫 모델(Model)을 선보였다. 이러한 성과는 앞에서 말한 몽골어 대사전 5권을 수정·보완하여 웹 사전으로 전환했기에 가능했던 것이다. 그리고 2015년에는 국내외에 적극적으로 사전에 대한 홍보 활동을 하고, 웹 디자인도 고쳤다. 같은 해 7월과 10월에는 단어 하나하나에 발음 기능과 표제어 몽골 옛 문자 표기 및 사진과 기타 자료 등을 추가하여 사전의 완성도를 높였고 그 해 12월 28일에 사업 결과를 발표했다. 그리고 2016년에는 모바일 앱 출시에 성공했고, 2017년에는 국내외로 홍보와 관련 학술 활동을 전개했다.

**그림 1** 개방형 몽골어 웹 대사전 표지



예를 들어, 검색창에 'aav(아버지)'라고 치면 '아버지'라는 단어 발음 기능과 함께 몽골 옛 문자 표기, 그리고 3개의 뜻풀이와 예문이 등장한다. 그 밑에 16개의 관련 연어와 관용어, 속담 표현도 해석과 함께 나온다(관련어 이하는 지면 관계상 생략함).

ААВ ᠠᠭᠪᠢ

1. Өөрийн эцгийг дуудах үг;
2. Төрсөн эцэг – Нарс гэдэг модоор наадмыг минь сийлсэн аав минь дээ хө. ААЗ, аав ээж[хоршоо] (эцэг эх) – Эцэг тэнгэр дэлгэ рэнгүй... В.Инжаннаши. Хөх судар., \*аавын чээж (а. Хаахгар, дав лиун зантай хүүхэд буюу залуу – Ахмад хүний дэргэд их үг хэлж болохгүй. Тэгвэл аавын чээж гэж муу хэлнэ. С.Дашдооров. Говийн өндөр; б.[яриа] УАЗ маркийн бага оврын ачааны автомаши

н), \*аавын чээж гаргах (нас бага хүүхэд хаахгар, давлиун зан гаргах) – Одоогийн багачууд даан ч настай хүний үг авахгүй, аавын чээж гаргаад байх болчихжээ, Д.Мягмар. Нэгэн байшингий н түүх., \*аавын шийр хатаах (аавынхаа ажил үйлсийг үргэлжлүү лэх), \*аавдаа адуу манахыг заах ~ аавдаа авгай авахыг заах (а. Мэдээжийн юмыг заах; б. Дэмий, илүү үг ярих), \*аавынд орох ~ \*аавынд суух ~ \*аавынд очих ~ аавын хаалга татах (шоронд орох), аавын бийд хүн тань, агтны бийд газар үз ~ аавын бийд хүнтэй танилц, агтны бийд газар үз [зүйр цэцэн үг] (аливаа ажил үйлийг тусалж зөвлөх хүн буй бололцоотой үед нь амжиж хий гэсэн санаа), аавын сургаал алт, ээжийн сургаал эрдэнэ [зүйр цэцэн үг] (аав ээжийн үг сургамж алт эрдэнээс үнэ цэнэтэ й гэсэн санаа), аавын хүү адилгүй, арван хуруу тэгшгүй [зүйр цэцэн үг] (хүн бүхэн адилгүй, хүний авьяас чадвар ч юм бүхэнд адилгүй гэсэн санаа);

3. Төрсөн эцэгтэй адилтган үзэх хүндтэй хүн: өргөсөн аав (үрчлэн авч өсгөсөн аав), тэжээсэн аав (өөрийн хүүхэд адил дэлгэрэнгүй...

출처: <https://monggollolli.mn/dictionary/>

현재 개방형 몽골어 웹 사전의 민간 참여도 현황을 살펴보면 아래와 같다.

- 민간 참여 인원: 1,395명(누적 수)
- 민간 참여 추가 표제어: 1,417개
- 민간 참여 추가 표제어 해석: 900여 개

이 웹 사전 편찬과 관련해서 조금 더 구체적인 정보를 얻기 위하여 사업 팀원이며 몽골과학아카데미 어문연구소 연구원인 푸레브수렝(T. Purevsuren) 박사를 인터뷰했다. 인터뷰 내용 중 현재 웹 사전 접속자 수와 웹 사전 차후 발전 계획에 대한 내용을 정리하면 아래와 같다.



- 웹 사전 접속자 수: 총 771,495명, 그중 615,375명(79.8%)이 새로운 접속자.
- 전 세계 119개국 2,646개 도시에서 접속(미국, 대한민국, 일본, 러시아, 중국, 독일 등 다른 국가에 거주하는 몽골인들이나 몽골에 관심이 있는 외국인으로 추정).
- 접속자들의 58%는 컴퓨터, 40%는 휴대 전화, 나머지 2%는 태블릿(tablet)을 활용.
- 10만 명이 모바일 앱을 통해 내려받기 한 상태이다.

(2018년 2월 28일 기준)

#### 개방형 몽골어 웹 사전의 차후 발전 계획

- 1) 사전 응용을 계속해서 확산하되 사전 개선 계획을 수립해 나간다.
- 2) 아직은 결정되지 않았지만 몇 년 주기로 종이형 대사전 출판을 고려하고 있다.
- 3) 민원 내용을 살펴보면 사전 사용자들이 단어의 기본 형태를 착각하거나 정서법 규정을 몰라 입력을 잘못해 단어가 검색되지 않는 경우에 불만을 표출하는데 앞으로 조금 틀리게 입력해도 찾고자 하는 단어가 검색될 수 있도록 기술적 보안을 하도록 한다.
- 4) 현재 제시된 용례가 완성된 문장 형태가 아닌 줄임 형태의 문장이 많아서 정확한 뜻풀이에 어려움이 생기곤 하는데 “The best of Mongolian literature electronic database and software of the 108 Volumes”라는 자체 개발 중인 프로그램을 통해 용례를 완성된 문장으로 바꾸는 작업을 시작하였으며 앞으로 웹 사전의 모든 용례를 수정·보완하도록 한다.
- 5) 현재 5,000개 정도의 기본 전문용어들이 수록되어 있는데 앞으로 방언형과 사용 빈도가 낮은 단어, 전문용어 수를 늘려서 웹 사전의 어휘 수를 늘린다.

#### 4. 차후 과제

이 글에서는 개방형 몽골어 웹 사전의 역사적 배경과 편찬 경과를 간단하게 살펴봤는데, 이는 한국의 개방형 사전 《우리말샘》처럼 편찬이 완전히 마무리되지 않은 사전이다(이현주, 2016:66). 아직 마무리되지 않았다는 것은 계속 변화와 발전 가능성이 있다는 의미이다.

따라서 몽골어 웹 사전의 경우 사전 개선 작업을 위해서는 몇 가지 과제가 남아 있다고 본다. 우선 민간 참여에서 전문가 투입이 좀 더 체계적으로 이뤄질 필요가 있다. 즉, 전문가 참여를 유도하기 위해 지원 및 격려 제도도 마련되어야 한다는 것이다.

그리고 웹 사전의 개선 문제와 지속적인 관리를 위해 인력 확보와 예산 지원이 필요하다. 지금은 몽골과학아카데미 연구원들의 일상적인 업무 틀에서 크게 벗어나지 않고 있는 실정이다. 물론 다양한 관련 사업 추진을 통해 어느 정도 개선될 여지는 있으나 더욱 체계적인 지원이 필요하다.

4차 산업 혁명을 논하고 있는 현 시점에서 전 세계 여러 나라에서 개발 중인 개방형 웹 사전은 몽골 국내뿐만 아니라 국제적인 확산도 고려하면서 다른 나라 전문가들과의 협력 사업 개발을 모색해야 한다. 더군다나 웹 사전인 만큼 정보 기술적인 발전과 성과를 적극 도입하여 빅데이터 및 맞춤형 활용도를 높이면 지금의 표제어 수를 훨씬 능가하는 훌륭한 웹 사전으로 발전하리라 믿어 의심치 않는다.

## 참고 문헌

- 배연경(2016), “사전의 진화와 미래”, 《새국어생활》 16-2, 국립국어원.
- 이현주(2016), “《우리말샘》 편찬 경과”, 《새국어생활》 16-4, 국립국어원.
- Л. Болд(2008), 《Монгол хэлний дэлгэрэнгүй тайлбар толь》,  
Оршил хэсэг, Монгол Улсын Шинжлэх Ухааны Академи,  
Улаанбаатар хот.
- Б. Ринчен(1964), 《Монгол бичгийн хэлний зүй》, Удиртгал,  
Улаанбаатар, 124-р тал.
- Ц. Шагдарсүрэн(2017), 《Монгол толь бичгийн уламжлал(То  
вч түүхэн тойм, анхааруштай зүйл)》, Эрдэм шинжилг  
ээний өгүүллийн чуулган - I боть “Монгол хэл шинжлэ  
л, алтай судлал”, 143-213 тал.
- Я. Цэвэл(1966), 《Монгол хэлний товч тайлбар толь》, Улаанба  
атар хот.
- Монгол Улсын Шинжлэх Ухааны Академи(2008), 《Монгол  
хэлний дэлгэрэнгүй тайлбар толь(Таван боть)》, Улаанб  
аатар хот.
- <https://mongoltoli.mn/dictionary/> Монгол Улсын Шинжлэх Ухаан  
ы Академи(2016), 《Монгол хэлний их тайлбар толь》.  
[ulemj.org/page/content/221](http://ulemj.org/page/content/221). Г. Мэнд-Ооёо(2013), 《Монгол хэл  
бичгийн сартваахи Шагж》.



국립국어원 소식

