

한국 방언 검색 프로그램 개발

이동광, 안동언, 정성중, *두길수, 김호영
전북대학교 컴퓨터공학과, *서남대학교 전기전자멀티미디어공학부

Implementation of Korean dialect Retrieval Program

Dong-Kwang Lee, Dong-Un Ahn, Sung-Jong Chung, Gil-Su Doo, Ho-Young Kim
Dept. of Computer Engineering, Chonbuk National University
Faculty of Electric, Electronic & Multimedia, Seonam University

E-mail : dklee@uan.chonbuk.ac.kr

요약

한국 방언 검색 프로그램은 21세기 세종계획 프로젝트의 한민족 언어 정보화 분과에서 개발한 것으로 남한의 방언과 북한의 방언 및 해외 동포가 살고 있는 지역에서 사용되고 있는 한국어 방언을 컴퓨터로 검색할 수 있도록 한 것이다. 한국 방언 검색 프로그램은 검색된 방언들의 정보를 HTML 문서로 화면에 표시할 수 있도록 되어 있으며, 방언의 각 지역별 사용 분포를 지도를 통해 확인할 수 있도록 되어 있다. 본 논문에서는 한국 방언 검색 프로그램의 기능과 방언 지도의 도시를 위한 지도 도시 알고리즘에 대해서 논한다.

I. 서론

한국 방언 검색 프로그램은 남한의 방언과 북한의 방언 및 해외 동포가 살고 있는 지역에서 사용되고 있는 한국어 방언을 컴퓨터로 검색할 수 있도록 하여, 21세기 지식기반 사회에서 국민들의 의사소통을 신속·정확하게 하며 또한 통일 이후에도 우리 민족 구성원들간의 의사소통을 원만하게 함으로써 민족의 정보 공유를 유도하여 민족문화 창달에 기여함을 그 목적으로 한다[1][2].

한국 방언 검색 프로그램은 21세기 세종계획 프로젝트의 한민족 언어 정보화 분과에서 개발하였다. 2000

년에는 프로그램 개발을 위한 기초 조사와 연구를 시행하였으며 2001년에는 남한의 방언을 검색할 수 있는 시스템을 개발하였다. 2002년에는 북한의 방언을 검색할 수 있는 시스템이 추가될 것이며, 2003년에는 중국을 중심으로 해외의 우리 동포들이 사용하는 방언을 검색할 수 있는 시스템을 개발함으로써 한국 방언 검색 프로그램 개발을 완료할 것이다.

방언은 의사 소통의 장애를 가져오며, 의사 소통의 장애는 우리 나라 문화 발전에 장애를 가져오게 된다. 따라서 각 방언 사용자들의 원활한 의사 소통을 위해서는 각 방언들간의 차이와 대조 관계를 밝혀두어야 할 필요가 있다. 의사 소통의 원활성은 곧 정보 전달의 정확성과 신속성을 의미하게 되며, 더 나아가 문화 전달의 정확성과 신속성을 가져오며, 그것은 곧 문화의 발달로 이어진다[1].

이러한 필요성에도 불구하고 한국어 방언을 검색하는 프로그램은 그 동안 개발되지 않았다. 각 지역 방언에 대한 조사는 꽤 이루어졌지만 우리 방언 전반에 대한 전반적인 조사가 이루어지지 않았기 때문이다.

한국 방언 검색 프로그램의 데이터베이스를 구축하기 위해서 지금까지 조사 연구된 30여권의 방언 자료집들 중에서 전자 파일로 수집한 21권의 방언 자료를 일정한 형태로 가공하여 정리하였다. 2001년에는 남한의 방언을 중심으로 개발하는 목표를 가지고 있지만, 북한 및 중국의 방언까지 같이 수집하고 정리하였다. 2001년도에 개발된 한국 방언 검색 프로그램에 수록된

표준어 어휘 항목은 19,432개이며, 방언 어휘 항목은 157,876개이다. 또한, 방언 지도에 표시된 군 단위의 세부 지역은 259개이다.

본 논문에서는 한국 방언 검색 프로그램의 구성과 기능을 설명한다. 특히, 방언 지도를 표시하기 위하여 사용된 지도 도시 알고리즘에 대해서 설명하고 방언 도시를 위한 데이터베이스와 지역 코드의 구조에 대해서 논한다.

II. 한국 방언 검색 프로그램 설계

전국의 다양한 방언을 보다 쉽게 검색할 수 있도록 범위검색기능과 복수 검색기능을 제공하였고, 방언의 분포를 보다 쉽게 이해할 수 있도록 방언의 사용지역을 지도로 출력할 수 있도록 설계하였다.

2.1 시스템 구성

① 검색

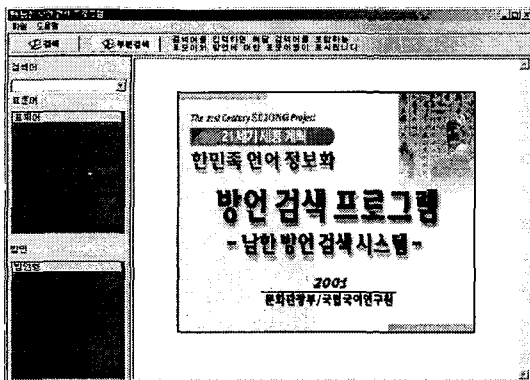
사용자가 입력한 검색어를 포함하는 표준어를 검색하고 그에 해당하는 방언을 화면에 출력해 준다.

② 부분검색

사용자가 검색할 범위나 복수 검색어를 입력받아 표준어를 검색하고 그에 해당하는 방언을 화면에 출력해 준다.

③ 지도

사용자가 선택한 방언이 사용되는 지역을 화면에 출력하여, 방언의 분포를 보다 쉽게 이해 할 수 있다.



<그림 1> 한국 방언 검색 프로그램 초기화면

한국 방언 검색 프로그램은 Microsoft Visual C++ 6.0을 사용하여 개발하였고, 데이터베이스는 Microsoft Access 97을 사용하였다. 본 시스템은 Windows98 이상 버전의 OS 환경에서 사용이 가능하다.

3.1 데이터베이스의 구성

한국 방언 검색 프로그램은 표준어 정보를 담고 있는 StandardDB 테이블, 방언 정보를 담고 있는 LocalDB 테이블, 지역별 코드 정보를 가지고 있는 MapCode 테이블로 구성된다.

[표 1] StandardDB 테이블의 구조

필드명	설명
ID	표준어 코드 번호, 원시 자료를 입력할 때 자동으로 생성한다.
Writer	작성자
Date	작성일
SWord	검색어, 표준어에서 공백 문자를 제거하여 검색에서 공백으로 인한 검색 실패를 없앤다.
Standard	표준어
Locals	표준어에 대한 방언형들의 목록
Pos	품사
Mean	표준어의 의미
Title	한자어나 고어가 포함된 표준어, Standard 에는 한자어나 고어가 포함되지 않는다.

[표 2] LocalDB 테이블의 구조

필드명	설명
ID	방언형 코드 번호, 원시 자료를 입력할 때 자동으로 생성한다.
Writer	작성자
Date	작성일
SWord	검색어, 방언형에서 공백 문자를 제거하여 검색에서 공백으로 인한 검색 실패를 없앤다.
Standard	방언형에 대한 표준어의 StandardDB 테이블의 ID를 지정한다.
Origin	원문과 방언형에 대한 참고 자료를 저장한다.
Map	방언형의 지역 분포
Example	방언형의 사용 예
Mean	방언형의 의미

III. 한국 방언 검색 프로그램의 구현

을 이용하여 색을 다르게 한다. <그림 4>는 이와 같은 마스크로 사용되는 이미지를 표시하고 있다.



(a) 경기도



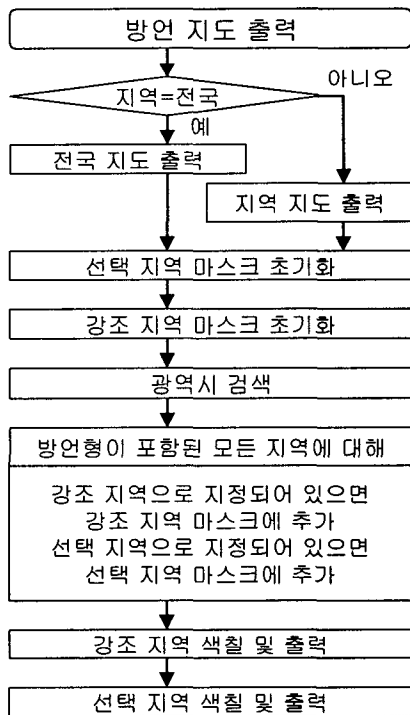
(b) 가평군



(c) 고양시

<그림 4> 지역 마스크

지도에서 단순히 목록에 있는 지역과 마우스로 클릭하여 선택한 지역은 서로 다른 색으로 표시되어야 한다. 이를 위해 바탕 지도에 지역 마스크를 덮어씌우기 전에 지역 마스크에 먼저 색을 칠한다.



<그림 5> 방언 지도 도시 흐름도

IV. 결론

본 논문에서는 한국 방언 검색 프로그램의 기능과 데이터베이스 구축에 대해서 설명하였다. 한국 방언 검색 프로그램에 사용된 방언 자료는 지금까지 조사 연구된 30여권의 방언 자료집들 중에서 전자 문서 파일로 수집한 21권의 방언 자료를 일정한 형태로 가공하여 정리한 것이다.

한국 방언의 지역 분포를 그림으로 확인할 수 있도록 방언 지도를 도시할 수 있도록 프로그램을 작성하였으며, 방언 지도의 행정 구역상의 표시를 위해 지역 코드를 사용하였다. 또한 지역 코드는 현재의 행정 구역 뿐만 아니라 옛 지명이나 읍·면 지역명을 사용할 수 있도록 확장 가능하게 설계했으며, 북한 및 중국과 일본 등에 거주하는 해외 동포가 사용하는 방언의 사용 지역도 표시할 수 있도록 설계하였다.

위와 같은 프로그램을 통하여 한국 방언의 분포를 쉽게 확인할 수 있으며, 방언 관련 연구에 기여할 것이다.

방언은 지리적, 역사적 배경에 따라 분포한다. 따라서 현재의 행정 구역 지도 이외에 산 및 강의 배치, 도로망 등의 배경 지도를 선택할 수 있는 연구가 필요하다. 또한, 역사에 따른 방언의 변천사를 연구하기 위하여 각 시대별 행정 지도에 방언 분포를 표시한다면 흥미로운 연구 결과를 얻을 수 있을 것이다.

V. 참고문헌

- [1] 이태영 외(2001), 21세기 세종계획 한민족 언어 정보화 연구보고서, 문화관광부/국립국어연구원
- [2] 세종계획 홈페이지, <http://www.sejong.or.kr/>
- [3] 김영태(1975), 경상남도 방언연구(I), 진명문화사
- [4] 이기갑 외(1997), 전남방언사전, 전라남도
- [5] 현평효 외(1995), 제주어사전, 제주도
- [6] 한국정신문화연구원(1987-1995), 한국방언자료집
- [7] 이상규(2000), 경북 방언사전, 태학사
- [8] 서울대학교(1997), 한국 방언사전
- [9] 표준국어대사전(국립국어연구원)에 등재되어 있는 방언 관련 자료
- [10] 우리말큰사전(한글학회)에 등재되어 있는 방언 관련 자료
- [11] 국어대사전(금성사)에 등재되어 있는 방언 관련 자료
- [12] 한영목(1999), 충남 방언의 연구와 자료, 이회문화사
- [13] 김병제(1980), 방언사전, 과학백과사전출판사
- [14] 김영배(1997), 평안방언연구(자료편), 태학사
- [15] 김태균(1986), 함북방언사전, 경기대학교 출판국