

CosmoScriBe 2.0: 한국어 전사 도구의 개발

CosmoScriBe 2.0 : The development of Korean transcription tools

곽선동* · 장문수**†

Sun-Dong Kwak, and Moon-Soo Chang†

*서경대학교 전자컴퓨터공학과, **서경대학교 컴퓨터공학과

† School of Computer science, Seokyeong University

요 약

구어 연구에서는 음성 데이터를 문자로 옮기는 전사(Transcription)라는 과정이 필요하다. 전사 작업을 보조하는 프로그램을 전사 도구라고 하는데, 발화 내용을 비롯하여, 발화 시간, 화자 정보 등의 많은 정보를 기록하는 다양한 기능을 제공한다. 이로 인하여 컴퓨터 사용에 익숙하지 않은 사용자는 숙지하는데 어려움이 있다. 또한 전사 도구는 국내에서 개발된 것이 거의 없어서 한국어 환경에 적합하지 않는 경우가 많다. 본 논문에서는 효율적인 한국어 전사를 지원하면서 비숙련자도 도구를 쉽고 빠르게 적용할 수 있는 전사 도구를 제안한다. 이를 위해 비숙련자를 위한 사용자 친화적인 인터페이스 환경을 제공한다. 또한 전사 과정에서 발생할 수 있는 실수를 최소화하기 위해 전사 지원 기능을 제공한다. 마지막으로 데이터 신뢰성을 위한 시스템 구조를 제공한다. 제안하는 도구에 대해 전사 경험의 유무에 따라 사용성 평가를 하였으며, 평가결과는 전체적으로 전사 속도 향상 및 전사 지원 기능이 편리한 것으로 나타났다.

키워드 : 전사 도구, 코퍼스 분석 도구, 마크업 언어, 사용자 인터페이스, 발화

Abstract

In spoken language research, transcription process needs to be carried out to translate voice data into text. Transcription tool, support program of transcription, offers various information such as content and time of utterance and speaker information. For this reason, inexperienced computer users are having trouble familiarizing with the program. Moreover, since there are little transcription tools developed domestically in Korea, they are usually not suitable for Korean environment. In this paper, we propose a transcription tool which supports not only Korean transcription but easy-to-use interface environment for novice. The transcription supporting function is also provided to minimize mistake that might happen in the process of transcription. And a system structure will be provided for data reliability. Usability of the proposed tool is evaluated in accordance with transcription experience. The evaluation result shows that transcription process and transcription support function have become faster and more convenient respectively.

Key Words : Transcription tool, Corpus Analysis Tools, Markup Language, User Interface, Utterance

1. 서 론

구어(spoken language)의 기본 데이터는 음성이지만, 분석을 위해서는 정보가 문자로 표현되어 있어야 한다. 따라서 구어자료로 수집한 음성 데이터를 문자로 옮기는 작업이

필요하다. 이 과정을 전사(Transcription)라고 한다. 최근에는 음성 인식 기술이 발달하여 자동으로 전사하는 경우도 있다. 그러나 구어연구 자료는 일상 대화, 아동의 대화, 전화 대화 등 여러 부류의 화자와 다양한 상황에서의 녹음 자료들이기 때문에 음성 인식 기술로는 모든 대화를 옮기는 것이 불가능하다. 따라서 기본적으로 전사는 사람이 수작업으로 한다.

전사는 소리로 구성된 정보뿐만 아니라 구어가 발생하는 시점, 말하는 대상 등의 주변 정보를 포함하여 대상의 말소리를 문자 형태로 기록하는 작업을 의미한다. 연구 주제에 따라 형태소 정보, 어절 정보, 레이블링¹⁾ 정보 등을 추가로 기록하는 경우도 있다. 각 정보는 티어(tier)²⁾라는 단위로

접수일자: 2013년 10월 2일

심사(수정)일자: 2013년 11월 11일

게재확정일자: 2013년 12월 16일

† Corresponding author

본 논문은 한국연구재단 연구자지원사업(NRF-2011-32A-B00202)에서 지원하여 연구하였음.

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

1) 전사 작업에서 레이블링이란 미디어 정보와 전사 자료를 연결해주는 작업을 의미한다. 즉 미디어의 타임라인 중 특정 시간에 발화 텍스트를 연결하는 작업이다.

기록된다.

초창기 수기(手記)로 시작한 전사는 컴퓨터의 발달로 타이핑 방식으로 변했다. 전사는 일반적인 텍스트 문서 작업과 다르게 녹음기, 비디오와 같은 미디어 도구를 함께 사용해야 한다. 따라서 미디어 도구와 함께 범용 텍스트 편집 프로그램으로 작업하면 프로그램 간에 자주 이동을 해야 하기 때문에 작업하기가 매우 불편하다. 최근에는 전사 전용 프로그램이 연구, 개발되고 있다[1][2]. 전사 전용 도구에는 EXMaRaLDA[3], Transcriber[2], CosmoScriBe 1.0[4] 등이 있다.

EXMaRaLDA는 티어 또는 발화자별 입력 기능이 제공된다. 음성 데이터 제어가 가능하고 전사에 필요한 다양한 기능들을 제공하고 있다. 특히 레이블링된 데이터의 분할, 통합이 가능하다. 한글은 정상적으로 저장은 되지만 도구에서 불러왔을 때 일부 컨트롤러에서 한글이 깨지는 현상이 있다. 이 도구는 다양한 기능을 제공하지만 인터페이스가 복잡하여 비숙련자가 사용하기에는 상당히 복잡한 구조로 되어 있다. Transcriber는 비교적 단순한 인터페이스를 사용하는 것이 특징이다. 레이블링과 데이터 입력이 동시에 이루어진다. 단순한 인터페이스는 비숙련자가 쉽게 도구에 적용할 수 있게 하지만 인터페이스의 제약이 많다. 예를 들어 오디오를 재생하지 않은 상태에서는 새로운 발화를 입력할 수 없다. 또한 레이블링의 위치 수정이 불가능하다. CosmoScriBe 1.0은 한 화면에서 비디오, 오디오의 재생과 텍스트 입력이 가능하다는 특징이 있다. 비디오와 오디오의 싱크 조절 기능, 미디어 속도 조절, 구간 반복 기능 등 미디어 컨트롤의 편리성에 특화되어 있다.

몇몇 전사 도구들은 티어별 입력 기능에 특화되어 있어 일반적인 텍스트 편집기의 인터페이스와는 다른 화면 구성으로 되어 있다. 따라서 전사를 처음 접하는 사용자 혹은 전사 경험이 있더라도 그 도구를 처음 접하는 사용자는 도구를 익숙하게 사용하기 위해 일정 시간의 학습이 필요하다. 또한 한국어 전사의 경우에는 해외 전사 도구를 사용하여 전사하는 경우도 있으나, 한글이 제대로 지원되지 않아 불편한 경우가 많아서 범용 텍스트 편집기와 녹음기를 이용한 작업이 일반적이다.

본 논문에서는 한글을 지원하고, 사용자가 쉽게 접할 수 있는 인터페이스를 제안하여 도구에 익숙해지는 시간을 최소화하고자 한다. 그리고 전사 작업에 유용한 텍스트 자동 완성, 헤더 자동완성 등의 전사 작업 지원 기능을 도입하여 편리한 전사 환경을 제공하고자 한다. 또한 자동 저장 및 마크업 데이터 구조를 제안하여 전사 과정에서 데이터 보호와 구축된 데이터의 신뢰성을 향상시킨다.

2. CosmoScriBe 2

본 논문에서 정의하는 사용자에게 편리한 도구란 작업 효율을 향상시키는 다양한 기능들을 구비하고, 이 기능들이 직관적인 인터페이스로 표현되어 쉽고 빠르게 적용할 수 있는 도구를 의미한다. 전사 도구에서 편리한 도구를 개발하기 위해 본 논문에서는 UI(User Interface), 전사 지원 기

능, 시스템 구조의 세 가지 측면에서 아이디어를 제안한다.

2.1 UI(User Interface)

인터페이스는 사용자와 도구가 가장 밀접하게 연관되는 영역이다. 인터페이스의 위치나 구조의 변화에 따라 사용자가 느끼는 사용감이 달라진다. 본 논문에서 제안하는 전사 도구의 이전 버전인 CosmoScriBe 1.0은 미디어 도구와 텍스트 편집기를 따로 동작시켜야 했던 과거의 전사 환경을 개선시키는 장점이 있었지만, 미디어 제어와 텍스트 입력 인터페이스를 한 화면에 구성함으로써 인하여 저해상도 환경에서는 화면 크기에 제약을 받는 문제점을 가지고 있었다. 전사 작업 특성상 연속적인 작업 도중에는 미디어 도구 화면보다 텍스트 편집 영역에 사용자의 시선이 고정된다. 따라서 미디어에 대한 시각적 정보, 즉 비디오 화면이나 오디오 컨트롤 등은 꼭 필요할 때 말고는 대부분의 작업 시간동안에는 화면에 나타낼 필요가 없다. 제안하는 전사 도구에서는 미디어 인터페이스와 텍스트 인터페이스를 분리하여 사용자의 선택에 따라 화면을 구성할 수 있도록 한다.

2.1.1 텍스트 입력 인터페이스

전사 환경에 오디오/비디오 등의 미디어 자료가 보편화되면서 기존 전사도구들은 전사 자료와 미디어 자료와의 링크를 중시하여 레이블링 작업을 중심으로 개발되었다. 이로 인해 발화 텍스트는 미디어의 타임라인에 부속되는 리스트 박스에 입력하는 방식이 채택되었다. 이 방식은 발화 텍스트를 미디어의 타임라인에 시각적으로 표시할 수 있는 장점이 있다. 그러나 텍스트가 레이블링 단위로 잘라서 입력되고 표시되기 때문에 초보자는 어느 정도 학습이 필요하다.

텍스트 편집 도구의 일반적인 환경은 텍스트 박스 방식이다. 전사 작업도 기본적으로 텍스트를 입력하고 편집하는 작업이므로 텍스트 박스 방식의 인터페이스가 사용자에게 더 익숙하다. 또한 발화 내용을 연속적으로 표기하기 때문에 주변 발화의 내용을 한눈에 확인할 수 있는 장점이 있다. 즉 주변 발화의 내용을 한눈에 확인할 수 있다.

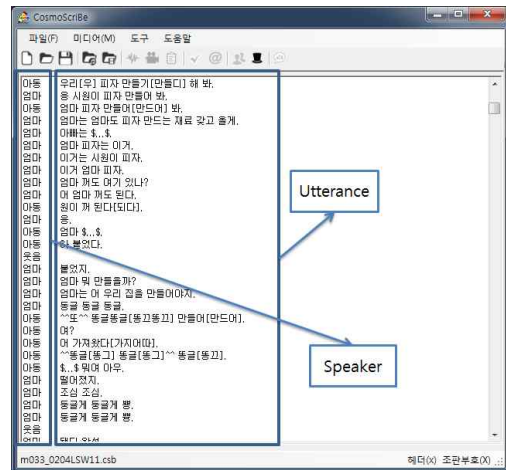


그림 1. 텍스트 입력 인터페이스
Fig. 1. Text input interface

본 논문에서는 사용자가 쉽게 접할 수 있는 텍스트 박스 방식의 입력 인터페이스를 사용하여 보다 쉽게 도구에 적용할 수 있도록 한다. 그림 1은 제안하는 입력 인터페이스를

2) 티어란 계층으로 전사의 경우 발화 티어, 레이블링 티어, 형태소 티어 등이 존재한다. 연구 목적에 따라 티어의 종류는 다양하게 존재한다.

적용한 도구의 화면으로, 메모장이나 워드와 유사한 편집 기능이 구현되어 있다.

2.1.2 미디어 제어 인터페이스

제안하는 전사도구에서는 미디어 컨트롤러를 텍스트 인터페이스와 분리하여 구성한다. 이 미디어 컨트롤러는 사용자 설정에 따라 활성화시킬 수 있어 화면 구성을 자유롭게 할 수 있다. 그림 2는 미디어 컨트롤러 창을 보여주고 있다.

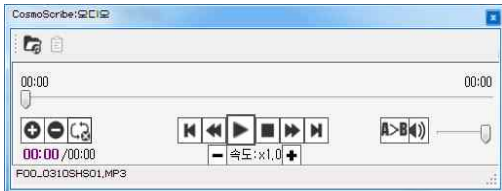


그림 2. 미디어 제어 인터페이스
Fig. 2. Media control interface

전사 작업은 귀로 음성을 들으면서 손으로는 키보드로 텍스트를 입력하는 매우 집중을 요구하는 작업이다. 작업하는 도중에 손이 키보드와 마우스를 번갈아 이동하게 되면 집중하는데 방해가 되어 작업 효율이 떨어지고 작업자의 피로감이 커지게 된다. 본 논문에서는 이것을 개선하기 위하여 전사 과정에서 주로 사용하는 미디어 제어 명령을 키보드 단축키로 구현하여 텍스트 인터페이스 상에서 미디어 제어가 가능하도록 한다. 표 1은 제안하는 미디어 제어 관련 단축키들을 나타낸다. 혼동되는 것을 방지하기 위하여 미디어 제어를 위한 모든 특수키 조합을 컨트롤키를 이용하도록 하였다.

표 1. 미디어 제어 단축키

Table 1. Hotkey for media control

Hotkey	Function
Ctrl + right, left	Move the playback position
Ctrl + up, down	Change the playback's speed
Ctrl + Space	Play / Pause
Ctrl + ;	Set up the repeat interval

2.2 전사 작업 지원 기능

전사 작업은 사람의 손으로 이루어지는 데이터 구축 과정이다. 따라서 작업 과정에서 발생할 수 있는 실수를 최소화하기 위한 기능이 필요하다. 또한 작업이 완료된 후에도 검토 작업이 필요하다. 본 절에서는 헤더 마법사, 자동입력 기능, 코퍼스 분석 도구를 통해 실수를 최소화하고 검토 작업을 간편하게 하는 전사 작업 지원 기능을 제안한다.

2.2.1 헤더 마법사

데이터 구축에는 부가적인 정보를 저장하기 위한 구조가 필요하다. 전사 자료에는 각 전사 파일에 헤더 정보로 기록하는 것이 일반적이다. 전사를 위한 헤더에는 화자 정보, 녹음 상황 정보, 전사 환경 정보 등이 기록된다.

헤더 정보는 말뭉치 수집 규칙에 따라 다르게 결정된다. 각 정보는 헤더 규칙(또는 전사 지침)에 따라 입력된다. 헤더 정보 중 발화자의 정보는 동일한 형식으로 여러 명의 정보를 입력하기 때문에 순차적으로 한 명씩 데이터를 입력하

는 것이 편리하다. 본 논문에서는 이러한 정보를 헤더 마법사 기능으로 구현한다. 헤더 마법사는 헤더에 입력될 정보를 사용자가 이해하기 쉽게 분할하여 순차적으로 입력하도록 한다. 그리고 사용자가 입력한 정보를 헤더 마법사가 헤더 규칙에 맞게 데이터를 생성한다. 본 논문에서는 KCLD 전사 지침³⁾을 기준으로 헤더 데이터를 생성한다. KCLD는 CHILDES[5]와 유사한 헤더 구조를 가지고 있다. 따라서 CHILDES를 활용하는 시스템과 호환이 편리하다.

```
@ Begin
@ Language: Kor
@ Participants: 1 - 아동, 2 - 엄마, 3 - 아빠, 4 - 이모
@ ID: koreansamples|CHI|3:3.||||Child|
@ ID: koreansamples|MOT|3:3.||||Mother|
@ Birth of CHI: 23-NOV-2007
@ Date: 04-JUL-2011, 11-JUL-2011
@ Transcriber: Min-sil Choi
@ Reviewer: Jae-eun Cha
@ Memo:
```

그림 3. 텍스트 형식의 헤더 정보
Fig. 3. Text header information

그림 3은 헤더 마법사를 통해 생성한 헤더 데이터를 나타낸 것이다. 이 정보를 사용자가 텍스트 에디터로 입력하게 되면 실수로 잘못 입력할 가능성이 매우 크다.

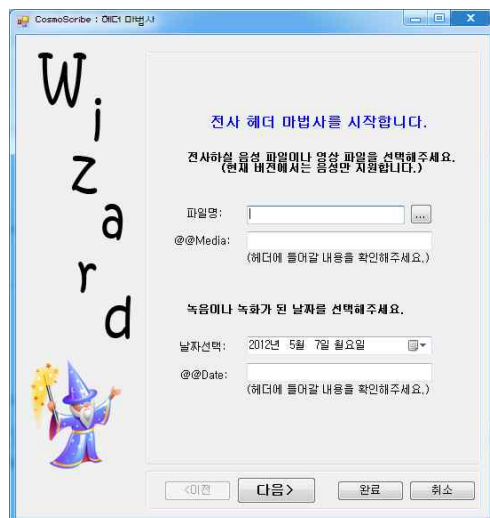


그림 4. 헤더 마법사
Fig. 4. Header wizard

그림 4는 헤더 마법사의 시작페이지를 나타내고 있으며, 화자 수에 따라 페이지 수가 다른데, 평균적으로 다섯 페이지 정도로 구성되어 있다. 콤보 박스, 라디오 버튼 등을 이용하여 사용자가 쉽고 정확하게 헤더 정보를 입력할 수 있

3) KCLD(Korea Child Language Database)는 한국어 아동이 일상생활에서 사용하는 어휘를 수집하여 데이터베이스화하는 한국연구재단의 기초연구과제의 결과물로 2011년도부터 수집이 진행되고 있는 어휘데이터베이스이다. KCLD 전사지침은 구축과정에서 전사를 위한 지침으로 규정된 것을 말한다.

도록 도와준다.

2.2.2 자동입력기능

전사 텍스트는 화자명과 발화 내용으로 구성된다(그림 1 참조). 전사자는 발화 음성을 듣고 말한 사람이 누구인지 파악하여 먼저 화자명을 쓰고, 들은 발화 내용을 그 옆에 입력하게 된다. 보통 화자명은 영문 코드로 되어 있다. 한국어 전사에서는 매번 한글/영어 전환키를 눌러야 하고, 같은 화자 코드를 각 발화마다 반복적으로 입력해야 한다. 그리고 화자명과 발화 내용 사이에는 탭(Tab)과 같은 구분자를 항상 입력해야 한다. 본 논문에서는 이러한 일련의 과정을 자동으로 처리하고 이 기능의 단축키로 숫자 키를 사용한다. 전사에서는 숫자도 한글로 입력하기 때문에 숫자 키를 발화 내용 입력에서는 사용하지 않으므로 혼란을 주지 않는 장점이 있다.

전사에는 언어 외적인 요소나 특수한 언어 표현 등을 나타내기 위하여 특별한 문자 조합 기호를 사용한다. KCLD 전사지침에는 예를 들어, '< >'는 의성어를 표기하는 기호이고, '[']'는 잘못 발화된 단어에 대한 원형복원 단어를 표기하는 기호이다. 한글을 입력하는 도중에 이 조합기호들을 사용하기 위해서는 복잡한 키 입력 과정이 필요하여 전사자들이 피로감을 호소하는 부분 중 하나이다. 또한 기호 입력 오류도 빈번하게 발생한다. 본 논문에서는 발화자 입력에 사용하고 남은 숫자 키(6~0)를 이용하여 문자 조합기호를 자동으로 생성한다. 표 2는 각 숫자 키에 할당된 자동 생성 기능을 나타내고 있다.

표 2. 단축키 기능
Table 2. Hotkey function

Hot Key	Function
1~5	Speaker + separator
6	< >, onomatopoeia
7	[], original restore
8	\$.\$. \$, undetermined
9	@!@, Comment
0	^^ ^^, repetition

2.2.3 코퍼스 분석 도구

전사에서 검토과정은 이차 전사라고 하며 녹음자료를 들으면서 전사된 텍스트를 다시 확인하는 작업이다. 이차 전사에서 띄어쓰기나 오타자와 같은 일부 오류들은 코퍼스 분석도구를 이용하여 수정할 수 있다. 음절 정보를 이용하여 '갈', '황'과 같은 한국어에 없는 음절이 포함되어 있는지 확인할 수 있다. 위와 같은 음절은 전사 과정에서 오타자일 확률이 높다. 또한, 어절 빈도는 “우리집에갔는데”와 같이 긴 어절이 검색된다면, 띄어쓰기가 잘못된 것을 알 수 있다. 그러나 코퍼스 분석 도구를 이용하려면 기존 코퍼스 분석 도구에서 요구하는 형식으로 데이터를 변환해야 한다. 검토 작업을 위해 매번 데이터를 변환하는 것은 사용자에게 매우 번거로운 일이다. 본 논문에서는 검토 작업에 필요한 코퍼스 분석 기능을 전사도구에 포함시켜, 보다 편리하게 이차 전사 작업을 할 수 있게 한다. 그림 5는 코퍼스 분석 기능을 추가함으로써 간략해진 검토 작업을 순서도로 표현한 것이다. 그림과 같이 별도의 파일 생성 및 저장 없이 검토를 할 수 있다.

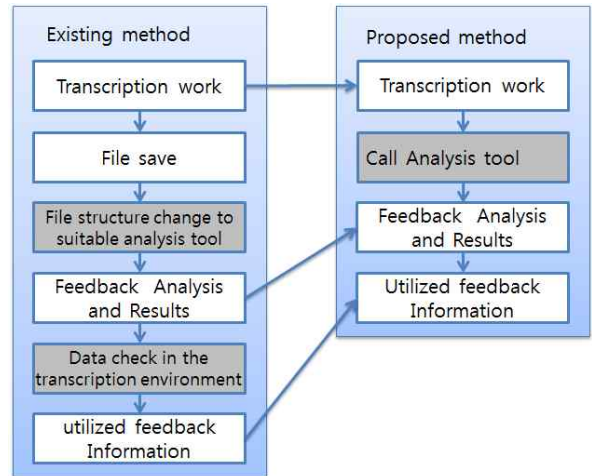


그림 5. 코퍼스 분석도구 사용 플로우
Fig. 5. Corpus Analysis Tool Flow

본 논문에서는 이차 전사를 보조하는 코퍼스 분석 기능으로 다음과 같은 기능들을 제안한다.

- 어절 및 음절의 빈도와 분포 통계
- 어절 및 음절 목록
- 키워드 검색
- 데이터 분할

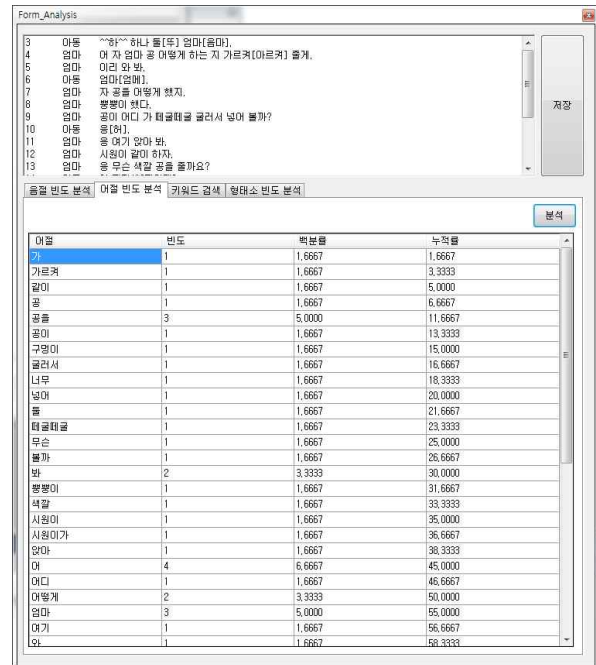


그림 6. 코퍼스 분석도구 : 어절 목록
Fig. 6. Corpus Analysis Tool : Eujeul List

어절 및 음절에 대한 통계 값과 목록은 위에서 언급한 것과 같이 오타자 확인에 유용하게 사용된다. 또한 이들에 대한 통계 값들은 구축된 전사 자료에 대한 통찰을 가능하게 하며, 향후 음성 수집에 참고가 될 뿐만 아니라 목적에 따라 전사 자료를 분류하거나 분할할 때 중요한 기준으로 사용된다. 그림 6은 어절에 따라 목록을 빈도순으로 나타내고

있는 화면이다. 화면 상단에는 대상 전사 텍스트를 보여주고 하단에 어절 목록을 보여준다. 목록의 각 칼럼을 기준으로 재배열이 가능하다.

코퍼스 분석 기능 중 키워드 검색은 발화 텍스트에서 특정 키워드가 출현하는 위치를 찾아주고, 해당 키워드의 앞, 뒤에 나타나는 어절 정보를 표시해준다. 이것은 이차 전사에서 특정 단어의 사용 예를 검색함으로써 일차 전사자가 올바르게 전사를 했는지 확인할 수 있게 해준다. 그리고 구축된 전사 코퍼스를 활용할 때 용례 검색 도구로 활용할 수 있다. 그림 7은 키워드 검색 도구의 일부를 나타낸 것이다. 검색과 출력 양식은 대표적인 한국어 코퍼스 분석 도구인 잠잠새[6]의 키워드 검색 방식을 따르고 있다.

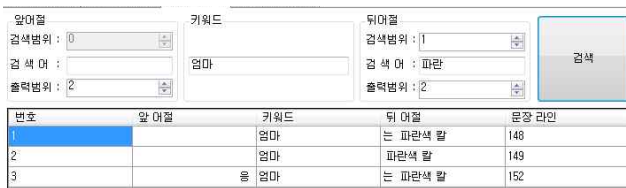


그림 7. 코퍼스 분석도구 : 검색 인터페이스
Fig. 7. Corpus Analysis Tools: Search Interface

코퍼스 분석 기능 중에서 마지막으로 제안하고 있는 데이터 분할 기능은 사용자 경험(UX)을 활용한 개발 방식을 통해 구현된 대표적인 기능으로써, 전사 과정 중에 필요에 따라 전사 텍스트 중 일부 영역을 잘라서 추출하는 기능이다. 추출 구간은 다음과 같이 사용자 선택에 의해 결정된다.

- 전체 발화 번호
- 화자별 발화
- 지정 구간 발화

전체 발화 번호로 추출하는 것은 모든 발화 텍스트에서 원하는 라인만큼 지정해서 추출하는 기능이다. 화자별 발화 추출은 지정한 화자의 발화 중 특정 개수를 추출하는 것으로, 그 사이에 있는 다른 발화자의 발화 수는 고려하지 않는다. 그리고 추출할 때 옵션에 따라 해당 화자의 발화만 추출하거나 그 사이에 있는 다른 화자의 발화도 포함할 수도 있다. 이 기능이 전사자들에게 매우 유용한 기능으로 평가되었다. 마지막으로 지정 구간 발화는 사용자가 발화 텍스트를 눈으로 확인하면서 마우스 드래그로 설정한 구간의 발화를 추출하는 기능이다. 그림 8은 추출하는 발화 구간을 나타내는 화면이다. 화자 선택 콤보 박스와 추출 구간의 시작과 끝 발화 번호가 나타나 있다.



그림 8. 데이터 분할 인터페이스
Fig. 8. Data partitioning interface

2.3 데이터 구조와 자동 저장

전사는 데이터를 구축하는 과정이므로 데이터의 보존과 신뢰성이 매우 중요하다. 전사 데이터는 헤더 정보, 미디어 정보, 발화 정보 등 다양한 정보를 포함하고 있어서 단일

파일에 기록하기 위해서는 구조화된 데이터 구조가 필요하다. 또한 저장된 데이터가 손상된 경우에 시스템에서 확인할 방법이 필요하다. 본 논문에서는 마크업 언어 형식의 데이터 구조를 사용하여 다양한 정보를 구조화하여 한 파일에 저장하고 저장된 데이터에서 태그 값을 이용해 오류가 있는 데이터인지 판단한다. 그림 9는 발화 정보의 일부를 마크업 언어 형식으로 표현한 데이터이다. 발화 단위의 태그와 화자명, 발화 정보를 태그와 속성값으로 표현한다.

```
<Utterance>
  <Speaker>엄마</Speaker>
  <Tier, type=text>이리 와 봐.</Tier>
</Utterance>
<Utterance>
  <Speaker>아동</Speaker>
  <Tier, type=text>엄마[엄메].</Tier>
</Utterance>
<Utterance>
  <Speaker>엄마</Speaker>
  <Tier, type=text>자 공을 어떻게 했지.</Tier>
</Utterance>
<Utterance>
  <Speaker>엄마</Speaker>
  <Tier, type=text>뽀뽀이 했다.</Tier>
</Utterance>
```

그림 9. 마크업 형식의 데이터 구조
Fig. 9. Mark-up types of data structures

최근의 문서 편집기는 자동 저장 기능이 포함되어 있어서 문서 작업 도중 비정상 종료가 일어나더라도 데이터를 보존할 수 있다. 본 논문에서는 문서편집기의 자동 저장과 유사한 기능을 구현하기 위하여 자동 저장 알고리즘을 개발한다.

자동 저장은 저장 간격이 짧을수록 데이터를 확실하게 보존할 수 있지만 시스템에 주는 부담이 증가한다. 따라서 시스템에 부담을 덜 주는 범위 내에서 저장 간격을 조절할 수 있는 알고리즘이 필요하다.

전사 작업은 화자와 발화 내용을 쌍으로 하여 한 줄씩 입력하게 된다. 본 논문에서는 화자와 발화 내용의 쌍을 데이터의 기본 단위로 설정하고, 이 기준을 자동 저장에도 적용한다. 즉, 하나의 발화를 입력, 수정, 삭제할 때마다 자동 저장이 이루어진다. 이러한 저장 빈도 조절을 통하여 큰 용량의 파일에서의 시스템 부하를 줄일 수 있다. 시스템 오류가 발생하게 되면 입력중인 발화의 데이터만 손실되고 나머지 데이터는 자동 저장된 데이터로부터 복구가 가능하게 된다.

3. 시스템 활용 및 평가

그림 10은 제안하는 도구를 사용하여 전사 작업을 진행할 때의 작업 흐름도이다. 작업자가 전사를 시작하면 먼저 새 문서를 생성하여 헤더 정보를 입력하고 미디어 파일을 불러온다. 텍스트 입력 컨트롤러에서 미디어 정보를 반복적으로 재생하면서 전사 텍스트 생성 작업을 수행한다. 전사 작업이 일정 수준이 되었을 때 분석 도구를 이용하여 1차 검토 작업을 한다. 1차 검토 작업에서는 오타자와 작성 중인 데이터가 원하는 데이터인지를 판단한다. 2차 검토 작업은 1차 전사가 완료된 데이터에 대해서 미디어 정보를 보면서 확인하는 과정이다.

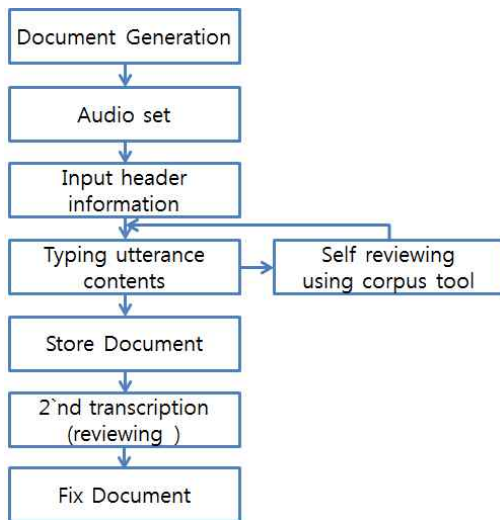


그림 10. 도구 사용 순서도
Fig. 10. Tools flowchart

본 논문은 비숙련자도 쉽게 적용할 수 있는 전사 도구를 제안하고 있다. 제안하는 시스템의 유용성 및 편리성을 검증하기 위하여 본 논문에서는 실제 전사 작업을 하는 작업자들을 대상으로 사용자 만족도 실험을 실시하였다. 실험에서는 21명의 작업자를 대상으로 하였으며, 전사 도구를 사용한 경험의 유무에 따라 두 그룹으로 나누어 실험을 실시하였다. 실험 방법은 제안하는 전사 도구를 일정 기간 동안 사용하게 하고 표 3과 같은 설문 내용에 대해서 답을 하도록 하였다. 응답은 1~5 사이의 값으로 표시되 만족도가 높을수록 큰 값을 주도록 하였다. 표 3에서 그룹 A는 6명이며, 실험 이전에 전사도구를 사용한 경험이 있는 사람들이다. 그룹 B는 15명으로 구성되어 있으며, 전사 도구를 처음 접하는 사람들로 구성되어 있다. 각 항목의 값은 응답의 평균값을 나타낸 것이다.

만족도 평균값은 3.87이며, 기존 전사환경에 익숙한 그룹 A보다 초보자인 그룹 B의 만족도가 높게 나타났다. 이것은 비숙련자를 고려한 설계의 효과인 것으로 판단된다. 특히 1번(안정성), 2번(프로그램 속도) 항목에서 두 그룹의 차이가 가장 크게 나타났다. 이것은 그룹 A가 개발 과정의 전사 도구 베타 버전을 사용했을 때 경험했던 오류가 영향을 준 것으로 판단된다. 속도에 대한 설문인 2번 항목에서 유일하게 보통 이하의 결과(2.73)가 나왔는데, 이것은 전사 데이터 용량이 커질수록 도구가 느려지는 현상 때문이었다. 이 오류는 설문 조사 후 알고리즘을 수정하여 개선되었다.

전사 작업 과정과 밀접한 관계가 있는 5번(데이터 입력), 6번(오디오 컨트롤), 10번(특수키) 항목에서 만족도가 매우 높게 나타나, 이 항목들의 만족도 평균값은 전체 평균보다 훨씬 높은 4.26으로 나타났다. 이 결과는 제안하는 도구의 편의성이 사용자들을 매우 만족시키고 있음을 반증한다. 이로 인해 두 그룹 모두 전사 작업 속도가 64%이상 향상된 것으로 나타났다.

4) 작업의 특성상 한 번 작업하는 데 걸리는 시간이 매우 길고, 한 사람이 같은 데이터로 두 번의 실험을 할 수 없기 때문에 주관적 평가 값을 이용하였다.

표 3. 평가표
Table 3. Checklist

	Contents	Group A	Group B
1	Is the program stable?	3.00	3.83
2	Is the program's speed fast enough?	2.73	3.50
3	Does the program speed up the work more than when the program is not being used?	3.87	4.33
4	Does the program make the work more accurate better than when it is not being used?	3.67	4.00
5	Is transcription data entry convenient?	4.13	4.00
6	Is media (audio) control convenient?	4.33	4.67
7	Are the automatic storage and the data management convenient?	3.87	4.17
8	Is the program easy to learn and use?	3.60	3.50
9	Is the program easy and convenient to use?	3.80	4.00
10	Is the shortcuts convenient to use?	4.13	4.33
	Total	3.71	4.03
		3.87	

5. 결론 및 향후 연구

본 논문에서는 편리한 한국어 전사 도구를 개발하면서 비숙련자도 쉽게 사용할 수 있는 전사 도구의 개발을 제안하고 있다. 이를 위하여 본 논문에서는 전사 도구를 사용자 친화적인 인터페이스, 전사 작업 지원 기능, 안정적인 시스템 구조와 같은 세 가지 영역에서 특화시켰다.

제안하는 도구는 사용자 만족도 평가에서 전사 경험 유무에 관계없이 작업 효율이 증가하고 만족도가 향상된 것으로 나타났다.

다만 안정성과 속도 면에서 평균치 이하의 만족도가 나타났는데, 특히 용량이 큰 전사파일을 도구에서 불러왔을 때 속도가 떨어지는 현상에 의한 결과였다. 평가 실험 후 알고리즘의 개선으로 상당 부분 개선이 이루어졌으며 향후에도 개선할 계획이다.

그리고 비숙련자를 위해 텍스트 편집이 편리한 텍스트 박스 형태의 인터페이스 방식을 채택하였는데, 이것은 레이블링 작업에서는 단점으로 작용할 수 있다. 향후 레이블링 기능을 추가할 경우 미디어 인터페이스 부분을 보완하여 이 문제점을 개선할 계획이다.

References

[1] Hetherington, L. & McCandless, M. "SAPPHIRE: an extensible speech analysis and recognition tool based on Tcl/Tk," *In Proceedings of ICSLP'96*, pp. 1942-1945, 1996.

- [2] Claude Barras, Edouard Geoffrois, Zhibiao Wu, and Mark Liberman, "Transcriber: development and use of a tool for assisting speech corpora production," *Speech Communication - Special issue on speech annotation and corpus tools*, Vol. 33, No. 1-2, pp. 5-22, 2001.
- [3] Thomas Schmidt, "EXMARaLDA and the FOLK tools - two toolsets for transcribing and annotating spoken language," *Proceedings of LREC*, pp. 236-240, 2012.
- [4] Moon-Soo Chang, "Korean transcription support tools developed Ver.1.0", *Proceedings of the 2nd KASA & KSHA joint conference*, O-2, pp.32-34, 2011.
- [5] Brian MacWhinney, "The CHILDES Project - Tools for analyzing talk, Part 1: The CHAT Transcription Format," Available: <http://childes.psy.cmu.edu/manuals/CHAT.pdf>, 2013, [Accessed: September 10, 2013].
- [6] Kim nam-don, "A Study on devices of Korean grammar in premodern era through using both the corpus of new novels and Kamchaksae program," *Korean language and literature*, Vol. 143, pp.123-157, 2006.

저 자 소 개



곽선동(Sun-Dong Kwak)
 2011년 : 서경대학교 소프트웨어학과 졸업
 2011년~현재 : 서경대학교 대학원
 전자컴퓨터공학과 석사과정
 관심분야 : UI, Image Processing
 Phone : +82-10-6329-9929
 E-mail : sigred@naver.com



장문수(Moon-Soo Chang)
 1992년 : 고려대학교 전자전산공학과 공학사
 1994년 : 고려대학교 전자공학과 공학석사
 2001년 : 동경공업대학 지능시스템과학전공
 공학박사
 2000년~2003년 : 한국전자통신연구원
 선임연구원
 2003년~현재 : 서경대학교 컴퓨터과학과 부교수

 관심분야 : Natural Language Understanding, Knowledge Mining, HCI
 Phone : +82-2-940-7754
 E-mail : cosmos@skuniv.ac.kr