

한글 철자 오류 교정 시스템

채영숙, 이영식, 권혁철

부산대학교 전자계산학과

A Correcting method of Speller for Hanguk

Chae Young-Soog, Lee Young-Sik, Hyuk-Chul Kwon

Dept. of Computer Science, Pusan National University

요약

문서 속에서 나타나는 철자 오류는 맞춤법을 몰라서 잘못 쓴 경우와 자료 입력 과정에서 잘못 입력된 경우로 볼 수 있다. 이들 각각의 오류들에 대해 맞춤법 및 표준어 규칙을 비롯하여 한국어 자료의 통계적 분석을 통한 교정 기법을 제공하고자 한다.

본 논문은 철자 검사기가 틀리다고 판단한 오류 어절로부터 올바른 어절을 찾아서 교정해 주고 철자가 틀린 원인을 알려줄 수 있는 도움말 기능을 제공하는 철자 교정 시스템을 구현하고자 한다.

I. 서론

사무 자동화의 시대가 시작되면서 컴퓨터를 이용한 문서 작성, 편집 및 전자 출판 문화가 보편화 되고 있다. 미국이나 일본의 경우는 단순한 문서 작성의 기능을 가진 1 세대 워드 프로세서에서 이제는 CRITIQUE와 같은 문서 편집을 비롯하여 글자 오류, 문법 오류까지도 교정해 주는 제 3 세대 워드프로세서가 실용화 되고 있다. 이러한 시대적 추세에 힘입어 한글용 워드프로세서에도 철자 검색 기능을 추가하고자 하는 노력들이 생기는 것을 볼 수 있다.

한국 과학 기술원의 HSPELL, 서울대학교와 IBM의 맞춤법 검사 교정기, 금성사의 철자 검사 기능을 추가한 문서 편집기 등 많은 대학 및 기업체에서 현재 연구 중에 있다. 하지만 교착어라는 한국어의 언어적 특성상 문서상에 나타난 맞춤법 오류는 물론이고 문법적인 오류를 검색하고 교정한다는 것은 쉬운 작업이 아니다.

본 연구실에서는 실용화를 목표로 한 한글 맞춤법 검색 및 교정 시스템(한글 철자 검사기)을 연구 개발 중에 있다. 본 논문에서는 한글 철자 검색기에 의해 발견된 틀린 어절을 대상으로

* 본 논문은 한국통신 소프트웨어 연구소 연구비 지원에 의한 결과임.

하며, 사용자가 이 오류 어절을 교정하는데 필요한 후보 어절 생성 방법, 즉 교정 기법을 비롯하여 맞춤법 지식과 같은 도움말을 제공하는 시스템을 소개하고자 한다. 이 시스템은 자체 개발한 문서 편집기와 자동 코드 변환기를 가지고 있으며, 현재 IBM PC상에서 수행된다.

II. 맞춤법 및 표준어 규정

(1) 맞춤법 규정

"맞춤법이란 그냥 글자를 맞추어서 글을 쓰는 방법을 이르는 것이 아니라, 사회적인 제도로써 정해진 규범에 따라 바르게 적어 나타내는 법, 곧 정서법을 이르는 것이다. 그러므로 그 언어 공동체의 구성원은 누구나 이를 따르고 지켜야 할 의무를 지닌다."라고 맞춤법을 정의하기도 한다.

현재 사용되는 맞춤법은 여러 번의 개정을 거쳐 1988년 1월 14일 문교부 고시 제 88-1호로 1989년 3월 1일부터 시행되고 있다. 이 규정은 크게 총칙을 비롯하여 소리, 형태, 띄어쓰기에 관한 규정으로 구성되어 있다. 이러한 규정들 중 많이 틀리는 대표적인 규정 예를 들어 보면 다음과 같다.

[규정 예 1 : 제 8 항]
 '계, 례, 매, 폐, 해'의 '해'는 '해'로 소리나는 경우가 있더라도 '해'로 적는다.

이 규정에는 예외로 사용되는 어휘들이 있다. 예를 들면 '휴계실'과 같은 경우는 '휴계실'이 틀린 표기법이고 '휴계실'이 맞는 표기법이다. 이처럼 예외 어휘들로 인해 혼란을 불러 일으킨다.

[규정 예 2 : 제 18 항]
 다음과 같은 용언들은 어미가 바뀔 경우, 그 어간이나 어미가 원칙에 벗어나면 벗어나는 대로 적는다.
 (18-6) 어간의 끝 'ㅂ'이 'ㄷ'로 바뀔 적
 가깝다 : 가까워 가까우니 가까웠다

위의 ㅂ 불규칙 규정은 " '듭-', '-곱'과 같은 단음절 어간에 어미 '-아'가 결합되는 경우는 '와'로 적는다"는 예외 규정이 있어 ㅂ 불규칙 사용('와/-워')에 혼란을 불러 일으킨다.

[규정 예 3 : 제 30 항]
 사이시옷은 다음과 같은 경우에 받치어 적는다.
 1. 순 우리말로 된 합성어로서 앞말이 모음으로 끝난 경우
 2. 순 우리말과 한자어로 된 합성어로서 앞말이 모음으로 끝난 경우

위의 사이시옷 규칙에도 한자어에 대해서 부분적인 예외가 있다. '셋방, 숫자, 횡수, 곳간, 찻간, 뿔간'의 6 단어에만 사이시옷을 결합해 허용할 뿐 한자만의 결합에는 사이시옷을 사용하지 않는다는 규정에 틀린 경우를 볼 수 있다.

[규정 예 4 : 제 40 항]

어간의 끝 음절 '하'의 'ㅏ'가 줄고 '하'이 다음 음절의 첫소리와 어울려 거센 소리로 될 적에는 거센 소리로 적는다.

위의 경우 안을림 받침 뒤에서는 '하'가 줄 적에는 준 대로 적는다는 조항에 의해 '깨끗지'가 아니라 '깨끗지'가 옳다는 사실을 몰라 잘못 쓰는 경우도 나타난다.

[규정 예 5 : 제 41 항 ~ 제 50 항]

- (1) 조사는 그 앞말에 붙여 쓴다.
- (2) 외존 명사는 띄어 쓴다.
- (3) 단위를 나타내는 명사는 띄어 쓴다.
- (4) 수를 적을 적에는 '만' 단위로 띄어 쓴다.
- (5) 두 말을 이어 주거나 열거할 적에 쓰이는 말들은 띄어 쓴다.

위의 띄어쓰기 규정은 문서상에서 가장 많은 빈도로 나타나는 오류의 유형이다. 이는 띄어쓰기 규정에 관해 확실한 지식을 가지고 있지 않으면 올바른 표기를 하기가 힘든 규정 중에 하나로 보인다.

[규정 예 6 : 제 54 항]

다음과 같은 접미사는 원소리로 적는다.

- (1) '-군/꾼'은 '꾼'으로 통일하여 적는다.
- (2) '-갈/깎'은 '깎'로 통일하여 적는다.
- (3) '-대기/매기'는 '매기'로 통일하여 적는다.

위의 접미사 외에도 '-배기/뺨기', '-적다/찍다'는 발음되는 유형에 따라 혹은 의미에 따라 달리 표기함으로 사용하는 이의 혼란을 불러 일으킨다.

(2) 표준어 규정

"표준어란, 한 나라 안에서 지역적, 사회적(계층적, 집단적) 차이를 초월한다. 가장 바람직한 의사 전달의 수단으로써 통일되고 규범화된 말인데, 이는 국가를 배경으로 하여 한 나라의 공식 용어로 사용된다."라고 표준어 규정을 정의하기도 한다.

현재 사용되는 표준어 규정은 1988년 1월 14일 문교부 고시 제 88-2호로 1989년 3월 1일부터 시행되고 있다. 이 규정은 크게 표준어 사정 규칙과 표준 발음법으로 구성된다. 하지만 이 중 교정 기법에 필요한 부분은 표준어 사정 원칙이다. 표준 발음법은 교정 기법 중의 하나인 음소 변환 규칙에 이용될 한국어 자료의 일부분이 된다. 표준어 사정 원칙은 총칙과 발음 변화에 따른 표준어 규정, 어휘 선택의 변화에 따른 표준어 규정으로 구성되어 있다. 이러한 규정들 중 가장 많이 틀리는 대표적인 예를 들면 다음과 같다.

[규정 예 7 : 제 10 항]
 다음 단어는 모음이 단순화한 형태를 표준어로 삼는다. (ㄱ을 표준어로 삼고, ㄴ을 버림)
 ㄱ. 괴팍하다. ㄴ. 괴팍하다 / 괴팍하다

위의 규정은 표준어 규정 중 가장 많은 오류가 나타나는 항목 중에 하나이다. 아직도 옛날의 규정에 익숙해 있는 사람은, 아니 규정이 바뀐 사실을 잘 모르는 사람은 이 규정을 자주 어긴다. 위의 예와 마찬가지로 복모음화 형태로 (케케묵다 -> 케케묵다) 사용하는 이가 많다.

[규정 예 8 : 제 17 항]
 비슷한 발음의 몇 형태가 쓰일 경우, 그 의미에 아무런 차이가 없고, 그 중 하나가 더 널리 쓰이면, 그 한 형태만을 표준어로 삼는다. (ㄱ을 표준어로 삼고, ㄴ을 버림)
 ㄱ. 꼭두각시 ㄴ. 꼭독각시
 내송스럽다 내홍스럽다

[규정 예 9 : 제 20 항]
 사어가 되어 쓰이지 않게 된 단어는 고어로 처리하고, 현재 널리 사용되는 단어를 표준어로 삼는다. (ㄱ을 표준어로 삼고, ㄴ을 버림)
 ㄱ. 오얏 ㄴ. 자두

위의 표준어 규정에 나타난 예 외에도 비슷한 경우의 오류들이 흔히 나타난다. 이러한 오류들은 자신이 사용되는 오류가 어떠한 이유에서 틀렸는 지조차도 모르는 경우가 허다하다. 따라서 우리 나라와 같이 맞춤법과 표준어 규정이 까다로운 나라에서는 사용자에게 자신이 틀린 이유를 설명해 줄 수 있는 도움말 기능이 요구된다.

III. 철자 오류의 유형에 따른 교정 기법

문서에 나타나는 철자 오류는 크게 두 가지 유형으로 분류할 수 있다. 하나는 문서를 작성하는 이가 맞춤법의 지식이나 표준말 사용이 익숙지 않아 발생하는 오류이다. 띄어쓰기 오류가 그 대표적 예이다. 다른 하나는 문서를 입력하는 과정에서 발생하는 타자수의 오류가 될 수 있다. 이러한 오류는 키보드 조작의 실수에 의해 발생하는 오류로 일정한 유형을 가지지는 않는다.

III.1. 맞춤법 오류

(1) 띄어쓰기 오류

띄어쓰기 오류는 가장 빈번히 나타나는 오류 유형이다. 특히 신문에서 자주 나타나는 오류로 단순히 지면을 절약한다는 관점에서 띄어쓰기 규정을 무시하지만 신문에 익숙해 있는 독자들의 문서에도 띄어쓰기 규정이 어긋난 경우들을 흔히 볼 수 있다. II의 [규정 예 5]와 같은 띄어쓰기 규정을 정확히 알고 사용하는 이가 드물다. 또한 띄어쓰기 규정은 품사나 사용 어휘의 의미에 따라 결정지어진다. 이러한 띄어쓰기 오류는 붙여 쓴 오류, 띄어 쓴 오류, 이 둘이 결합된 형태의 오류가 있다. 이들 오류 중 다음과 같은 알고리즘 기법을 이용해 후보 어절을 찾는다.

I. (붙여 쓴 오류)

```
fill-buffer ( string )
i <- string length
while (string[i] is not blank) {
  insert-blank ( string )
  check-speller ( string )
  if (correct)
    make replaced string
}
```

II. (띄어 쓴 오류)

```
fill-buffer ( string )
concatenate-buffer ( string, pre-string )
check-speller ( string )
if (correct)
  make replaced string
concatenate-buffer ( string, after-string )
check-speller ( string )
if (correct)
  make replaced string
```

III. 두 가지 결합 형태의 오류

```
ready concatenate-buffer ( string, pre-string )
ready concatenate-buffer ( string, after-string)
check-speller ( strings )
if (success in I)
  make replaced string
```

(2) 그 밖의 맞춤법 오류

띄어쓰기 규칙을 제외한 나머지 규정들은 ‘한글 맞춤법 통일안’ 규정을 최대한 준수한다. 앞 뒤 음절의 음소 비교에 의해 적합한 규칙들을 적용한다.

뿐만 아니라 각 규칙들에 알맞은 도움말 제시 기능이 있다. 제시되는 도움말은 맞춤법 규정이나 오류에 대한 교정 정보로서 사용자가 왜 틀렸는 지의 이유를 모를 때 참조할 수 있다.

비교 2음절	대치 2음절	대치 조건	맞춤법 도움말
N1 N2	R1 R2	C1	사이시옷 규칙오류
N1 N2	R1 R2	C2	불규칙 규칙 오류
N1 N2	R1 R2	C3	두음 법칙 오류

(Ni : 입력 문자열, Ri : 대치 문자열, Ci : 규칙 적용 조건)

< 표 1 : 맞춤법 적용 규칙 구조 >

III.2. 입력 오류

오류의 종류 중에서 편집자의 입력 오류는 그 오류의 유형이 일정하지 않다. 한글의 음소별 모든 치환이나 변경에 의한 대치 어절 생성은 시스템의 성능 저하 뿐만 아니라 편집자에게 엉뚱한 대치 음절을 제시하여 혼란을 더해 줄 수 있다. 그러므로 이의 처리는 많은 실험 데이터의 분석에 의한 통계적 처리를 우선한다. 또한, 편집자가 숙련된 타자수라면 사용하는 키보드의 자판배열 상태에 유의하여 음소의 변환 규칙을 정한다.

다음 <표2>는 철자검사기에서 쓰이는 사전 내의 단어들에 대한 음소 발생 빈도(%)이다. 본 철자 검사기에 사용된 사전의 품사 중 용언은 규칙 동사인 경우는 어미 ‘-다’를 뺀 나머지, 불규칙 용언인 경우는 불규칙 종류에 따라 사전 표제어를 달리 한다. <표2>의 용언 사전은 한국어의 허형 변화 특징상 올바른 수치라 볼 수는 없다. 그러나 명사 사전의 경우는 이 수치와 단어 평균 길이가 2.74 정도라는 점을 이용해 4번째 음절까지로 음소 변환을 제한해 교정 시스템의 성능 향상면에 이점이 된다.

사전	명 사 사 전			용 언 사 전		
	초성	중성	종성	초성	중성	종성
1	ㅇ: 15.4	ㅏ: 20.3	ㅜ: 42.6	ㅇ: 13.4	ㅏ: 21.3	ㅜ: 45.7
	ㄱ: 14.5	ㅓ: 14.3	ㄴ: 15.3	ㄱ: 12.4	ㅓ: 15.5	ㄴ: 16.0
	ㅈ: 12.3	ㅈ: 12.0	ㅇ: 15.3	ㅂ: 8.9	ㅈ: 13.0	ㅇ: 9.2
2	ㅅ: 13.7	ㅏ: 20.1	ㅜ: 39.8	ㅇ: 13.8	ㅏ: 22.9	ㅜ: 52.1
	ㅇ: 12.9	ㅣ: 14.1	ㅇ: 15.6	ㄱ: 11.7	ㅓ: 15.4	ㄱ: 11.5
	ㄱ: 12.6	ㅓ: 11.6	ㄴ: 14.2	ㄴ: 10.9	ㅣ: 15.0	ㅇ: 9.7
3	ㅈ: 15.7	ㅣ: 19.7	ㅜ: 46.7	ㄷ: 18.4	ㅏ: 37.4	ㅜ: 67.6
	ㅅ: 13.5	ㅏ: 19.4	ㅇ: 12.8	ㄱ: 12.5	ㅓ: 19.3	ㅇ: 8.8
	ㄱ: 13.0	ㅓ: 14.5	ㄱ: 10.8	ㅎ: 11.6	ㅣ: 12.9	ㄷ: 7.3
4	ㅇ: 18.2	ㅣ: 25.7	ㅜ: 53.9	ㄴ: 32.8	ㅣ: 37.2	ㅜ: 85.2
	ㄴ: 12.7	ㅏ: 16.6	ㄱ: 11.2	ㅎ: 20.0	ㅏ: 27.7	ㅇ: 3.4
	ㅈ: 11.3	ㅡ: 11.0	ㄴ: 10.8	ㄱ: 9.1	ㅓ: 11.8	ㄴ: 2.8

<표 2 : 음절별 음소 발생 빈도 수>

자음의 경우는 한국어 통계 분석에 따라 특히 대표음 법칙에 의해 틀린 어절에 음소 변환 규칙을 적용해 후보 어절을 찾는다. 모음의 경우는 단모음과 복모음의 혼란에서 오는 경우가 많다.

초성	{ 1번째 음절 } { n번째 음절 }				
	1순위	2순위	3순위	
ㄱ	ㄲ	ㅇ	ㅈ		
ㄴ	ㄷ	ㅇ	ㅉ		

< 표 3 : 음소 변환 규칙 >

III.3. 표준말 오류

이 오류는 지방적 특색을 강하게 나타내며, 규정의 변경에서 오는 혼란이 심하게 나타나는 규정이다. 표준어 오류는 많은 문서 속에서 발견된 오용어 사례들을 수집함으로써 교정이 가능하다. 이 규칙은 다른 규칙과는 달리 사용자 스스로가 틀린 이유를 모르는 경우가 많은 오류이다. 따라서 상세한 도움말 기능이 절실히 요구되는 부분이기도 하다.

III.4 휴리스틱 정보를 통한 후보 어절 제시

한글 철자 교정기는 사전에 의존한 구조이다. 따라서 신조어나 사회적인 변화를 나타내는 어휘들은 철자 검사기에서 틀렸다고 판단한다. 예를 들면 "국민된 도리"라는 예문에서처럼 "국민"은 "되다"라는 말과는 결합되어 쓰이지 않다가 '국민의 자격, 신분을 가진'이라는 의미로 그 뜻을 인정받아 현재 사용되고 있다.

이런 식의 사전 미등록어를 무조건 틀렸다고 판단하는 것보다는 '명사 + 용언' 형태라는 특정 유형의 오류에 대해서만 필요한 형식 형태소를 삽입한 후보 어절을 사용자에게 제시하고자 한다. 이러한 규칙들은 명사에 결합되는 용언의 종류에 따라 중간 삽입 형태소가 다르다. 예를 들면 "국민된"은 "국민이 된"으로 "-이"라는 주격 조사를 삽입한 형태이고 "학생같다"는 "학생과 같다"로 접속 조사를 추가한 형태이다.

	주격	목적격	접속(으로)	접속(과)
가다	X	X	0	X
같다	X	X	X	0
당하다	X	0	X	X
되다	0	X	X	X
받다	X	0	X	X
하다	X	0	X	X

< 표 4 : 매개 형식 형태소 삽입 규칙 >

III. 시스템의 설계 및 결과 분석

이상의 교정 규칙에 의해 철자 교정기를 검증하면 다음과 같다. 실험에 사용한 자료는 본 연구실에서 현재 보관 중인 언론 연구원의 4 대 주간지 신문, 부산일보에서 제공한 연합 통신 자료, 타자 1급 자격증을 가진 사람이 입력한 중학교 국어 교과서, 맞춤법의 지식이 부족한 사람이 쓴 소설 등이다.

	교 정 기 법					
	띄어쓰기	음소변환	삽입규칙	맞춤법	표준어	미교정
문서1	73	2	2	20	0	3
문서2	30	14	1	11	0	3
문서3	182	40	10	160	15	6

(문서 1 : 신문 문서 2 : 중학교 교과서, 문서 3 : 소설)

< 표 5 : 실험 결과 >

위 실험 결과치는 문서 특성에 따라 나타나는 오류의 종류가 다르다. 이들 오류 중 문서 전반적으로 많이 나타나는 유형은 띄어쓰기 오류이다.

IV. 결론

IBM-PC상에서 실행되는 본 연구실의 철자 교정기는 C-언어를 사용하여 구현했다. 한글 철자 검사기의 한 부분인 철자 교정기는 사용자의 문서 교정의 번거로움을 덜어 주기 위한 목적을 가진다. 뿐만 아니라 바뀐 맞춤법에 익숙지 않은 사용자에게는 틀린 어절에 대한 도움말을 제공함으로써 맞춤법 사용에 도움을 준다.

한국 언론 연구원과 부산 일보사의 신문, 중학교 국어 교과서 및 소설 등 한국어 자료로부터 오류 어절의 유형을 추출하여 분석했다. 이를 통해 얻어진 띄어쓰기 규칙, 음소 변환 규칙, 형식 형태소 삽입 규칙, 맞춤법 규칙 및 자주 틀리는 유형이나 표준어의 대치 규칙 등을 설정하여 구현하였다. 이러한 규칙들은 많은 자료 분석을 통해 얻어지는 규칙들로 확장, 변경의 편의성과 용이성을 고려하여 구현했다.

현재 본 연구실에서는 다량의 신문 자료 분석을 통해 한국어의 특성 및 교정에 필요한 규칙을 확장하고 있으며, 어절 단위의 교정 뿐만 아니라 어절 단위를 넘어선 교정까지도 사용자에 게 제공함으로써 컴퓨터를 이용한 문서 작성에 편의를 주고자 한다.

철자 검색

<한글 두벌식>H 148줄 30칸 삽입 E:\MAKEHSP\DATA

구체적으로 어떻게 하겠다는 것인지 불분명하나, 그 방향만은 본란에서

부사의 끝 음절이 분명히 '이'로만 나는 것은 '-이'로 적고, '히'로만 나가나 '이'나 '히'로 나는 것은 '-히'로 적는다.

(1) '이'로만 나는 것
(예) 가뿔이 깨끗이 나뿔이 느긋이

(2) '히'로만 나는 것
(예) 극히 급히 딱히 속히

(3) '이,히'도 나는 것
(예) 솔직히 가만히 간편히 나흔히

자음

ESC(Exit)

※ 또한 철자를 대략 자음에 발진하고 하더라도 공동편이 무엇이 어떨까 한다. 현실적으로 상당수 대학이 독자적인 편이다. 그런 대학을 포함, 희망하는 대학은 선택적으로 공동 편의 취지인 자율정신에도 부합된다. 경우에 따라서는 대학의 등장도 기대해볼 수 있을 것이다. ※ ◎15일

철자 검색

<한글 두벌식>H 79줄 63칸 삽입 E:\MAKEHSP\DATA

개혁방안을 연구하고 있다. 또 한 번의 학력고사, 한 걸음 결정한다는 것도 시정을 해야 되겠고, 입시과목이 너무 줄이고 학생들의 부담도 줄이는 방향으로 교육개혁을 대통령 연두연설 요지※ 올해는 민주주의를 정착시키는 데 있어, 또한 줄기 대로 도약하는 데 있어 가장 큰 고비가 되는 제도 통일의 길로 나아가는 결정적인 전기를 맞는 해가 될 수 있는 상황에서도 법과 질서, 안정의 바탕을 굳건히 세워

※ 30년 만에 다시 시행하는 지방자치는 이름다음에 피어나는 민주주의와 입헌군주주의의 결합이다. 지방자치의 성공을 위해서는 유권자인 국민이 적극적으로 참여하여야 합니다. ※ 정치권을 빌미로 스스로의 이익을 도모하려는 동태의 발전을 위해 헌신적으로 일할 일꾼을 뽑아주어야 합니다. ※ 경제의 안정은 우리 경제의 앞날을 좌우하는 결정적인 요인이 될 것이다. 이 또한 급속히 오를 경우 그나마 되살아나고 있는 회복 불능의 상태에 빠질 것이며 우리 경제도 저는 오늘 이 자리에서 근로자와 기업, 모든 경제주체가 인식하여 우리 경제의 안정과 발전을 위해

[참고 문헌]

- [1] N.H.Macdonald, L.T.Frase, P.Gingrich, and S, A.Keenan, "The WRITER'S WORKBENCH : Computer aids for text analysis", IEEE Trans. Commun. COMM-30, No.1, pp.105-110, 1982
- [2] G.E.Heidorn, Jensen, L.A.Miller, R.J.Byrd, and M.S.Chodorow, "The EPISTLE Text-Critiquing System", IBM Syst.J. Vol, No.3, pp.305-326, 1982
- [3] Stephen D.Richardson, "Enhanced Text Critiquing using a Natural Language Parser", research report RC-11332, IBM Thomas J.Watson Research Center, 1985
- [4] 이은정, 최신 표준어.맞춤법 사전, 국어 문화사, 1991.
- [5] 김창근, 한글 새 표기법, 도서출판 글숲, 1991.
- [6] 채영숙, "한글 철자 검색기와 교정기의 구현 개발 환경", 부산대학교 석사학위 논문, 1991
- [7] 김재원, "한글 맞춤법 오류의 교정 기법에 관한 연구", 부산대학교 석사학위 논문, 1992
- [8] 최기선 외 2인, "한국어 형태소 처리와 사전 -접속 정보를 이용한 한글 철자 및 띄어쓰기 검사기 ", 어학연구, PP.87-116, 1990.
- [9] 박종만, "철자 검색기에서 틀린 어절의 처리", 우리말 정보화 잔치 '91 논문집, 국어정보학회, pp.187-195, 1991